

STIC-Biotech/ChemLib

119041

From: Davis, Minh-Tam  
Sent: Thursday, April 08, 2004 10:34 AM  
To: STIC-Biotech/ChemLib  
Subject: Search request for 09/743825

Please search in commercial database, issued patent files, PGPUB and interference:

SEQ ID NO:1, 7, 8, 10

Thank you.

MINH TAM DAVIS

ART UNIT 1642, ROOM 3A24, MB 3C18

272-0830

RECEIVED  
APR - 8 2004  
STIC

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 4/18/04  
Date Completed: 4/15/04  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH: 4  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: 06p  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 119041

TO: Minh-Tam Davis  
Location: rem/3a24/3c18  
Art Unit: 1642  
Thursday, April 15, 2004

Case Serial Number: 09/743825

From: Edward Hart  
Location: Biotech-Chem Library  
REM-1A55  
Phone: 571-272-2512

edward.hart@uspto.gov

### Search Notes

Examiner Davis,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

*my lab*  
*07/98*

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 14:29:57 ; Search time 8.27166 Seconds  
(without alignments)  
9524.167 Million cell updates/sec

Title: US-09-743-825-8

Perfect score: 21

Sequence: 1 ctggcgatctctgaagagctctg 21

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2475585 seqs, 1875730760 residues

Total number of hits satisfying chosen parameters: 4951170

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	100.0	25	14	US-10-098-263B-19933
C 2	19.4	92.4	25	14	US-10-098-263B-19934
C 3	16.8	80.0	804	15	US-10-369-493-46860
C 4	16.8	80.0	227931	15	US-10-085-117-274
C 5	16.4	78.1	736	15	US-10-027-632-11566
C 6	16.4	78.1	2751	12	US-10-282-132A-40306
C 7	16.4	78.1	3530	15	US-10-108-260A-521
C 8	16.4	78.1	13435	10	US-09-764-891-6875
C 9	16.4	78.1	13435	15	US-10-242-355-1017
C 10	16.2	77.1	230	9	US-09-728-445-219
C 11	16.2	77.1	376	9	US-09-864-761-5179
C 12	16.2	77.1	414	9	US-09-815-242-9106
C 13	16.2	77.1	414	9	US-09-815-242-9430
C 14	16.2	77.1	414	12	US-10-282-132A-37546
C 15	16.2	77.1	439	9	US-09-764-870-80

C 16	16.2	77.1	439	9	US-09-764-853-36
C 17	16.2	77.1	439	9	US-09-764-860-223
C 18	16.2	77.1	439	14	US-10-125-540-80
C 19	16.2	77.1	439	14	US-10-074-095-223
C 20	16.2	77.1	439	14	US-10-103-313-14
C 21	16.2	77.1	439	15	US-10-212-872-223
C 22	16.2	77.1	439	15	US-10-158-057-13
C 23	16.2	77.1	584	15	US-10-027-632-48523
C 24	16.2	77.1	584	15	US-10-027-632-48524
C 25	16.2	77.1	682	15	US-10-027-632-289222
C 26	16.2	77.1	1096	12	US-10-425-114-13742
C 27	16.2	77.1	1450	12	US-10-425-114-36311
C 28	16.2	77.1	1457	12	US-10-425-114-34239
C 29	16.2	77.1	1587	12	US-10-282-132A-13145
C 30	16.2	77.1	1726	15	US-10-027-632-99323
C 31	16.2	77.1	1727	15	US-10-027-632-97269
C 32	16.2	77.1	1727	15	US-10-027-632-97270
C 33	16.2	77.1	1731	12	US-10-425-114-395
C 34	16.2	77.1	1977	12	US-10-425-114-36425
C 35	16.2	77.1	2811	12	US-10-425-114-27734
C 36	16.2	77.1	4423	9	US-09-764-870-578
C 37	16.2	77.1	4423	9	US-09-764-853-853
C 38	16.2	77.1	4423	9	US-09-764-860-1031
C 39	16.2	77.1	4423	14	US-10-125-540-578
C 40	16.2	77.1	4423	14	US-10-074-095-1031
C 41	16.2	77.1	4423	14	US-10-103-313-581
C 42	16.2	77.1	4423	15	US-10-212-872-1031
C 43	16.2	77.1	4423	15	US-10-158-057-341
C 44	16.2	77.1	10503	9	US-09-764-870-579
C 45	16.2	77.1	10503	9	US-09-764-853-854

ALIGNMENTS

RESULT 1

US-10-098-263B-19933/c  
; Sequence 19933, Application US/10098263B  
; Publication No. US20030104410A1  
; GENERAL INFORMATION:  
; APPLICANT: Mittman, Michael  
; TITLE OF INVENTION: Human Microarray  
; FILE REFERENCE: 3118.1  
; CURRENT APPLICATION NUMBER: US/10/098,263B  
; PRIOR FILING DATE: 2003-01-08  
; PRIOR APPLICATION NUMBER: 60/276,759  
; PRIOR FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 131066  
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1  
; SEQ ID NO 19933  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-098-263B-19933

Query Match 100.0%; Score 21; DB 14; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
|||||  
DB 24 CTGGCGTATCTGAAGAGTCTG 4

RESULT 2

US-10-098-263B-19934/c  
; Sequence 19934, Application US/10098263B  
; Publication No. US20030104410A1  
; GENERAL INFORMATION:  
; APPLICANT: Mittman, Michael  
; TITLE OF INVENTION: Human Microarray  
; FILE REFERENCE: 3118.1  
; CURRENT APPLICATION NUMBER: US/10/098,263B

```
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 19934
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-098-263B-19934

Query Match      92.4%; Score 19.4; DB 14; Length 25;
Best Local Similarity 95.2%; Pred. No. 2.8;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21
DB 24 CTGGCGTATCTGAAGAGTCTG 4

RESULT 3
US-10-369-493-46860/c
; Sequence 46860, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46860
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-10-369-493-46860

Query Match      80.0%; Score 16.8; DB 15; Length 804;
Best Local Similarity 90.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGGCGTATCTGAAGAGTCTG 21
DB 698 TGGCGTATCTGAAGAGTCTG 679

RESULT 4
US-10-085-117-274
; Sequence 274, Application US/10085117
; Publication No. US2003023334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 227931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: variation
; LOCATION: (1)...(227931)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-274

Query Match      80.0%; Score 16.8; DB 15; Length 227931;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGGCGTATCTGAAGAGTCTG 21
DB 147307 TGGGTATCTGAAGACTCTG 147326

RESULT 5
US-10-027-632-11566
; Sequence 11566, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11566
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-11566

Query Match      78.1%; Score 16.4; DB 15; Length 736;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GGCCTATCTGAAGAGTCT 20
DB 674 GGCCTATCTGAAGAGTCT 691

RESULT 6
US-10-282-122A-40306
; Sequence 40306, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
```

```
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40306
; LENGTH: 2751
; TYPE: DNA
; ORGANISM: Treponema pallidum
US-10-282-122A-40306

Query Match 78.1%; Score 16.4; DB 12; Length 2751;
Best Local Similarity 94.4%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGGTATCTGAAGAGTCT 20
DB 348 GCGGTATCTGAAGAGTCT 365

RESULT 7
US-10-108-260A-521
; Sequence 521, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 521
; LENGTH: 3530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-521

Query Match 78.1%; Score 16.4; DB 15; Length 3530;
Best Local Similarity 94.4%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGT 18
DB 1905 CTGGCTTATCTGAAGAGT 1922

RESULT 8
US-09-764-891-6875
; Sequence 6875, Application US/09764891
; Publication No. US20030077808A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6875
; LENGTH: 13435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-6875

Query Match 78.1%; Score 16.4; DB 10; Length 13435;
Best Local Similarity 94.4%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGT 18
DB 8309 CTGGCTTATCTGAAGAGT 8326

RESULT 9
US-10-242-355-1017
; Sequence 1017, Application US/10242355
; Publication No. US20030235831A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC003C1
; CURRENT APPLICATION NUMBER: US/10/242,355
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,897
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1017
; LENGTH: 13435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-355-1017

Query Match 78.1%; Score 16.4; DB 15; Length 13435;
Best Local Similarity 94.4%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGT 18
DB 8309 CTGGCTTATCTGAAGAGT 8326

RESULT 10
```

```

US-09-728-445-219
; Sequence 219, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020102543A1el Mutated Mammalian Cells and
; TITLE OF INVENTION: Animals
; FILE REFERENCE: LEX-0102-USA
; CURRENT APPLICATION NUMBER: US/09/728,445
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219
; LENGTH: 230
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(230)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-445-219

Query Match 77.1%; Score 16.2; DB 9; Length 230;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGAAGAGTCTG 21
    ||||| ||||| ||||| |||||
Db 184 CTGGTGTCTGAAGAGACTG 204

RESULT 11
US-09-864-761-5179
; Sequence 5179, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5179
; LENGTH: 376
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008496.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
US-09-864-761-5179

Query Match 77.1%; Score 16.2; DB 9; Length 376;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGAAGAGTCTG 21
    ||||| ||||| ||||| |||||
Db 200 CTGGCTGATCTCAAGAGTCTG 220

RESULT 12
US-09-815-242-9106
; Sequence 9106, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9106
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(414)

```

US-09-815-242-9106

Query Match 77.1%; Score 16.2; DB 9; Length 414;  
Best Local Similarity 85.7%; Pred. No. 1.8e+02;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
|||||  
Db 320 CTGGTGTATCTGAAGAGATTG 340

RESULT 13

US-09-815-242-9430  
Sequence 9430, Application US/09815242  
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
PROKARYOTES  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9430  
LENGTH: 414  
TYPE: DNA  
ORGANISM: Streptococcus pneumoniae  
NAME/KEY: CDS  
LOCATION: (1)...(414)  
US-09-815-242-9430

Query Match 77.1%; Score 16.2; DB 9; Length 414;  
Best Local Similarity 85.7%; Pred. No. 1.8e+02;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
|||||  
Db 320 CTGGTGTATCTGAAGAGATTG 340

RESULT 14

US-10-282-122A-37546  
Sequence 37546, Application US/10282122A  
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 37546  
LENGTH: 414  
TYPE: DNA  
ORGANISM: Streptococcus pneumoniae  
US-10-282-122A-37546

Query Match 77.1%; Score 16.2; DB 12; Length 414;  
Best Local Similarity 85.7%; Pred. No. 1.8e+02;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
|||||  
Db 320 CTGGTGTATCTGAAGAGATTG 340

RESULT 15

US-09-764-870-80/c  
Sequence 80, Application US/09764870  
Patent No. US20020042386A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT214  
CURRENT APPLICATION NUMBER: US/09/764,870  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 646  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 80  
LENGTH: 439  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (48)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (292)  
OTHER INFORMATION: n equals a,t,g, or c

```

; NAME/KEY: SITE
; LOCATION: (429)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-870-80

Query Match      77.1%; Score 16.2; DB 9; Length 439;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 CTGGCGTATCTGAGAGTCTG 21
        ||||| ||||| |||||
DB      230 CTGGCTGATCTCAGAGTCTG 210

```

Search completed: April 10, 2004, 20:21:40  
Job time : 9.27166 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:16:47 ; Search time 71.5529 Seconds  
(without alignments)  
8764.220 Million cell updates/sec

Title: US-09-743-825-8

Perfect score: 21

Sequence: 1 ctggcgatctgaagagtctg 21

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estum:*
16:	em_estom:*
17:	em_gss_hum:*
18:	em_gss_inv:*
19:	em_gss_pin:*
20:	em_gss_vrt:*
21:	em_gss_fun:*
22:	em_gss_nam:*
23:	em_gss_mus:*
24:	em_gss_pro:*
25:	em_gss_rod:*
26:	em_gss_phg:*
27:	em_gss_vrl:*
28:	gb_gsal:*
29:	gb_gsal2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	87.6	816	14	CD758735
C 2	17.8	84.8	370	12	B1566131
C 3	17.8	84.8	410	28	B2231447
C 4	17.8	84.8	459	12	B1566126

C 5	17.8	84.8	482	12	B1566371
C 6	17.8	84.8	508	13	BY469887
C 7	17.8	84.8	523	12	B1618797
C 8	17.8	84.8	542	12	B1566120
C 9	17.8	84.8	594	28	B2293209
C 10	17.8	84.8	631	13	BY179097
C 11	17.8	84.8	637	28	AQ787069
C 12	17.8	84.8	747	28	AQ780441
C 13	17.8	84.8	816	10	BG030798
C 14	17.8	84.8	846	13	BY730151
C 15	17.8	84.8	894	13	EX709246
C 16	17.4	82.9	463	28	AQ902638
C 17	17.4	82.9	503	28	AQ566793
C 18	17	81.0	250	28	B2175587
C 19	16.8	80.0	319	10	BF407422
C 20	16.8	80.0	319	14	W33528
C 21	16.8	80.0	379	28	AQ151626
C 22	16.8	80.0	399	9	AA966471
C 23	16.8	80.0	404	13	BY029119
C 24	16.8	80.0	426	9	AA906420
C 25	16.8	80.0	428	9	AA906422
C 26	16.8	80.0	451	13	BY483227
C 27	16.8	80.0	471	9	AA710085
C 28	16.8	80.0	490	10	AW767502
C 29	16.8	80.0	502	13	BQ261883
C 30	16.8	80.0	522	9	AA388360
C 31	16.8	80.0	529	28	BH222120
C 32	16.8	80.0	556	28	AZ877343
C 33	16.8	80.0	607	28	BH222188
C 34	16.8	80.0	616	12	B1751485
C 35	16.8	80.0	621	13	BQ258448
C 36	16.8	80.0	624	10	BQ663438
C 37	16.8	80.0	630	29	AG155633
C 38	16.8	80.0	639	28	AZ054680
C 39	16.8	80.0	653	29	CG115211
C 40	16.8	80.0	663	13	CA077550
C 41	16.8	80.0	664	14	CF665394
C 42	16.8	80.0	672	29	CG370684
C 43	16.8	80.0	673	29	CG370675
C 44	16.8	80.0	695	29	CG141551
C 45	16.8	80.0	727	29	CG441500

#### ALIGNMENTS

RESULT 1

CD758735

LOCUS

DEFINITION

IMAGE:695935 5', mRNA sequence.

CD758735

VERSION

KEYWORDS

SOURCE

ORGANISM

1 (bases 1 to 816)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CD758735 816 bp mRNA linear EST 30-JUN-2003  
AGENCY:14638907 NCI CGAP ZKId1 Danio rerio cDNA clone  
IMAGE:695935 5', mRNA sequence.

CD758735 GI:32343022

EST.

Danio rerio (zebrafish)

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Osteiophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 816)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgaabs-remail.nih.gov

Tissue Procurement: Leonard I. Zon, M.D.

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
 Plate: LLAM14592 row: k column: 06  
 High quality sequence start: 2  
 High quality sequence stop: 689.  
 Location/Qualifiers  
 1. 816

#### FEATURES

source  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7955"  
 /clone="IMAGE:6959935"  
 /lab\_host="DH10B (T1-resistant)"  
 /clone\_lib="NCI CGAP ZKId1"  
 /note="Organ: kidney; Vector: pCMV-SPORT6.1; Site:1:  
 EcoRV; Site:2: NotI; Cloned unidirectionally. Primer:  
 Oligo dt. Average insert size 1.8 kb. Constructed by J.  
 Wang (Research Genetics, Invitrogen Corp) from tissue  
 donated by L. Zon (Harvard University). Note: this is a  
 NCI CGAP Library."

ORIGIN  
 Query Match 87.6%; Score 18.4; DB 14; Length 816;  
 Best Local Similarity 95.0%; Pred. No. 4.9e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCT 20  
 |||||  
 DB 589 CTGGCGTATCTGATGAGTCT 608  
 |||||

RESULT 2  
 B1586131/c  
 LOCUS B1586131 370 bp mRNA linear EST 06-SEP-2001  
 DEFINITION RH25888.5prime RH Drosophila melanogaster normalized Head pF1c-1  
 Drosophila melanogaster cDNA clone RH25888 5 similar to EG:63B12.8:  
 FBan0014804 GO:[] located on: X 2A2-2A2:: 08/18/2001, mRNA  
 sequence.

ACCESSION B1586131 GI:15477553  
 VERSION B1586131.1  
 KEYWORDS EST.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
 1 (bases 1 to 370)  
 Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B.,  
 Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E.,  
 George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G.,  
 Misra, S., Mungall, C.J., Nuncio, J., Pacleb, J., Paragas, V., Park, S.,  
 Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and  
 Rubin, G.M.  
 BDGP/HMI RH Drosophila EST Project  
 Unpublished (2001)  
 Contact: Stapleton, M.  
 BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: http://www.fruitfly.org/EST\_est@fruitfly.berkeley.edu  
 hit genomic AB003422: arm:X [1512959,1813697]  
 estimated-cyto:1F4-2C7: 08/18/2001  
 Plate: RH.258 row: H column: 4  
 High quality sequence stop: 325.  
 Location/Qualifiers  
 1. 370  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"  
 /clone="RH25888"  
 /sex="male and female"  
 /dev\_stage="Adult"  
 /lab\_host="DHS-alpha Tona"

FEATURES  
 source  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /strain="BN/SENHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-380P3"  
 /sex="Female"  
 /cell\_type="Brain"  
 /clone\_lib="CHORI-230 Segment 2"  
 /note="Vector: pTARACL3; Site:1: MboI; Site:2: MboI;  
 CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by  
 Pieter de Jong"

ORIGIN  
 Query Match 84.8%; Score 17.8; DB 28; Length 410;  
 Best Local Similarity 90.5%; Pred. No. 7.3e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21

/clone\_lib="RH Drosophila melanogaster normalized Head  
 pF1c-1"  
 /note="Organ: head; Vector: pF1c1; Site:1: XhoI; Site:2:  
 BamHI; Library was kindly generated by Piero Carninci at  
 the RIKEN. The library was normalized and excised using  
 Cre recombinase. Plasmid cDNA library."

ORIGIN  
 Query Match 84.8%; Score 17.8; DB 12; Length 370;  
 Best Local Similarity 90.5%; Pred. No. 7e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
 |||||  
 DB 49 CTGGCGTATCTGTCGAGTCTG 29  
 |||||

RESULT 3  
 B2231447/c  
 LOCUS B2231447 410 bp DNA linear GSS 12-OCT-2002  
 DEFINITION CH230-380P3.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone  
 CH230-380P3, genomic survey sequence.

ACCESSION B2231447  
 VERSION B2231447.1 GI:23889988  
 KEYWORDS GSS.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE  
 1 (bases 1 to 410)  
 Zhao, S., Shetty, J., Shatsman, S., Tseng, G., Geer, K.,  
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,  
 Riggs, F., de Jong, P., and Fraser, C.M.  
 Rat BAC End Sequences from Library CHORI-230 MboI segment  
 Unpublished (1999)  
 Other GSSs: CH230-380P3.TJ  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/or ering information.htm). BAC end  
 page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 380 row: P column: 3  
 Seq primer: T7  
 Class: BAC ends.  
 Location/Qualifiers  
 1. 410  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /strain="BN/SENHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-380P3"  
 /sex="Female"  
 /cell\_type="Brain"  
 /clone\_lib="CHORI-230 Segment 2"  
 /note="Vector: pTARACL3; Site:1: MboI; Site:2: MboI;  
 CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by  
 Pieter de Jong"

FEATURES  
 source  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /strain="BN/SENHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-380P3"  
 /sex="Female"  
 /cell\_type="Brain"  
 /clone\_lib="CHORI-230 Segment 2"  
 /note="Vector: pTARACL3; Site:1: MboI; Site:2: MboI;  
 CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by  
 Pieter de Jong"

ORIGIN  
 Query Match 84.8%; Score 17.8; DB 28; Length 410;  
 Best Local Similarity 90.5%; Pred. No. 7.3e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21



Db 229 CTGGAGTATCTGAAGAGTCTG 209

||||| ||||| ||||| ||||| |||||

RESULT 4  
BI566126/c  
LOCUS  
DEFINITION  
RHS3268.5prime RH Drosophila melanogaster normalized Head pFlc-1  
Drosophila melanogaster cDNA clone RH3268.5 similar to EG:63B12.8:  
Fban0014804 GO:[ ] located on: X 2A2-2A2:: 07/31/2001, mRNA  
sequence.

ACCESSION  
BI566126  
VERSION  
BI566126.1 GI:15455318  
KEYWORDS  
EST.  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 459)

REFERENCE  
AUTHORS  
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,  
Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,  
George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,  
Misra,S., Mungall,C.J., Nunco,J., Pacleb,J., Paragas,V., Park,S.,  
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and  
Rubin,G.M.

TITLE  
BDGP/HMI RH Drosophila EST Project  
JOURNAL  
Unpublished (2001)  
COMMENT  
Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu  
hit genomic AE003422: arm:X [1512959,1813697]  
estimated-cyto:1F4-2C7: 07/31/2001  
Plate: RH.352 row: F column: 8  
High quality sequence stop: 458.  
Location/Qualifiers  
1. .459  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="RH35268"  
/sex="male and female"  
/dev\_stage="Adult"  
/lab\_host="DH5-alpha TonA"  
/clone\_lib="RH Drosophila melanogaster normalized Head  
pFlc-1"  
/notes="Organ: head; Vector: pFlc1; Site: 1: XhoI; Site 2:  
BamHI; Library was kindly generated by Piero Carninci at  
the RIKEN. The library was normalized and excised using  
Cre recombinase. Plasmid cDNA library."

FEATURES  
source  
Query Match 84.8%; Score 17.8; DB 12; Length 459;  
Best Local Similarity 90.5%; Pred. No. 7.6e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN  
1 CTGGCGTATCTGAAGAGTCTG 21  
||||| ||||| ||||| ||||| |||||  
Db 49 CTGGCGTATCTGCGAGTCTG 29

RESULT 5  
BI566371/c  
LOCUS  
DEFINITION  
RHS3593.5prime RH Drosophila melanogaster normalized Head pFlc-1  
Drosophila melanogaster cDNA clone RH3593.5 similar to EG:63B12.8:  
Fban0014804 GO:[ ] located on: X 2A2-2A2:: 08/01/2001, mRNA  
sequence.

ACCESSION  
BI566371  
VERSION  
BI566371.1 GI:15455563  
KEYWORDS  
EST.  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 459)

REFERENCE  
AUTHORS  
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,  
Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,  
George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,  
Misra,S., Mungall,C.J., Nunco,J., Pacleb,J., Paragas,V., Park,S.,  
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and  
Rubin,G.M.

TITLE  
BDGP/HMI RH Drosophila EST Project  
JOURNAL  
Unpublished (2001)  
COMMENT  
Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu  
hit genomic AE003422: arm:X [1512959,1813697]  
estimated-cyto:1F4-2C7: 07/31/2001  
Plate: RH.352 row: F column: 8  
High quality sequence stop: 458.  
Location/Qualifiers  
1. .459  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="RH35268"  
/sex="male and female"  
/dev\_stage="Adult"  
/lab\_host="DH5-alpha TonA"  
/clone\_lib="RH Drosophila melanogaster normalized Head  
pFlc-1"  
/notes="Organ: head; Vector: pFlc1; Site: 1: XhoI; Site 2:  
BamHI; Library was kindly generated by Piero Carninci at  
the RIKEN. The library was normalized and excised using  
Cre recombinase. Plasmid cDNA library."

FEATURES  
source  
Query Match 84.8%; Score 17.8; DB 12; Length 459;  
Best Local Similarity 90.5%; Pred. No. 7.6e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN  
1 CTGGCGTATCTGAAGAGTCTG 21  
||||| ||||| ||||| ||||| |||||  
Db 49 CTGGCGTATCTGCGAGTCTG 29

RESULT 6  
BY469887/c  
LOCUS  
DEFINITION  
RIKEN full-length enriched, melanocyte Mus musculus cDNA  
clone G270075A12 3', mRNA sequence.

ACCESSION  
BY469887  
VERSION  
BY469887.1 GI:26804266  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 508)

REFERENCE  
AUTHORS  
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,  
Nikaido,I., Ootani,N., Saito,R., Suzuki,H., Yamana,H.,  
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,  
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,  
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,  
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruscia,V.,  
Chochia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,  
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,  
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmer,S.,  
Grimmond,S.,

KEYWORDS  
EST.  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 482)

REFERENCE  
AUTHORS  
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,  
Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,  
George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,  
Misra,S., Mungall,C.J., Nunco,J., Pacleb,J., Paragas,V., Park,S.,  
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and  
Rubin,G.M.

TITLE  
BDGP/HMI RH Drosophila EST Project  
JOURNAL  
Unpublished (2001)  
COMMENT  
Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu  
hit genomic AE003422: arm:X [1512959,1813697]  
estimated-cyto:1F4-2C7: 08/01/2001  
Plate: RH.355 row: H column: 9  
High quality sequence stop: 457.  
Location/Qualifiers  
1. .482  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="RH35593"  
/sex="male and female"  
/dev\_stage="Adult"  
/lab\_host="DH5-alpha TonA"  
/clone\_lib="RH Drosophila melanogaster normalized Head  
pFlc-1"  
/notes="Organ: head; Vector: pFlc1; Site: 1: XhoI; Site 2:  
BamHI; Library was kindly generated by Piero Carninci at  
the RIKEN. The library was normalized and excised using  
Cre recombinase. Plasmid cDNA library."

FEATURES  
source  
Query Match 84.8%; Score 17.8; DB 12; Length 482;  
Best Local Similarity 90.5%; Pred. No. 7.7e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN  
1 CTGGCGTATCTGAAGAGTCTG 21  
||||| ||||| ||||| ||||| |||||  
Db 49 CTGGCGTATCTGCGAGTCTG 29

RESULT 7  
BY469887/c  
LOCUS  
DEFINITION  
RIKEN full-length enriched, melanocyte Mus musculus cDNA  
clone G270075A12 3', mRNA sequence.

ACCESSION  
BY469887  
VERSION  
BY469887.1 GI:26804266  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 508)

REFERENCE  
AUTHORS  
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,  
Nikaido,I., Ootani,N., Saito,R., Suzuki,H., Yamana,H.,  
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,  
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,  
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,  
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruscia,V.,  
Chochia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,  
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,  
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmer,S.,  
Grimmond,S.,

Gustringich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maita, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verdard, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E., and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
Aizawa, K., Kimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Cells were provided by Drs. William J Pavan, Stacie Loftus, and Denise Larson (Division of Intramural Research Genetic Disease Research Branch National Human Genome Research Institute, National Institutes of Health (NIH) Building: 49, Room 4A82 49 Convent Drive MSC 4472 Bethesda, Maryland U.S.A) whose assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

## SOURCE

1. 508  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="G270075A12"  
/cell\_type="melanocyte"  
/clone\_lib="RIKEN full-length enriched, melanocyte"

## ORIGIN

Query Match 84.8%; Score 17.8; DB 13; Length 508;  
Best Local Similarity 90.5%; Pred. No. 7.9e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGAAGAGTCTG 21  
||||| ||||| ||||| |||||  
Db 203 CTGGCGTATCTGAAGAGTCTG 183

## RESULT 7

BI618797/c

## LOCUS

DEFINITION BI618797 523 bp mRNA linear EST 07-SEP-2001  
RH49711.1 Prime RH Drosophila melanogaster normalized Head pFLC-1  
Drosophila melanogaster cDNA clone RH49711 5 similar to EG:63B12.8:  
Fban0014804 GO: {} located on: X 2A2-2A2;; 08/19/2001, mRNA  
sequence.

## ACCESSION

VERSION BI618797

## KEYWORDS

SOURCE EST.

## ORGANISM

Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

## REFERENCE

1 (bases 1 to 523)

## AUTHORS

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B.,  
Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E.,  
George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Ligo, G.,  
Misra, S., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,  
Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and  
Rubin, G. M.

## TITLE

BDGP/HMI RH Drosophila EST Project

## JOURNAL

COMMENT Unpublished (2001)

## CONTACT

Contact: Stapleton, M.

## BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)

hit genomic AB034422: armX (1512959,1813697)

estimated-cyto:IF4-2C7: 08/19/2001

Plate: RH.497 row: A column: 11

High quality sequence stop: 459.

## FEATURES

Source

1. 523

/organism="Drosophila melanogaster"

/mol\_type="mRNA"

/db\_xref="taxon:7227"

/clones="RH49711"

/sex="male and female"

/dev\_stage="Adult"

/lab\_host="DH5-alpha Tora"

/clone\_lib="RH Drosophila melanogaster normalized Head pFLC-1"

/note="Organ: head; Vector: pFLC1; Site 1: XhoI; Site 2: BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

## ORIGIN

Query Match 84.8%; Score 17.8; DB 12; Length 523;  
Best Local Similarity 90.5%; Pred. No. 7.9e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGAAGAGTCTG 21  
||||| ||||| ||||| |||||  
Db 49 CTGGCGTATCTGTCGAGTCTG 29

## RESULT 8

BI566120/c

## LOCUS

DEFINITION BI566120 542 bp mRNA linear EST 06-SEP-2001  
RH35246.5 Prime RH Drosophila melanogaster normalized Head pFLC-1

Drosophila melanogaster cDNA clone RH35246 5 similar to EG:63B12.8:  
Fban0014804 GO:[] located on: X 2A2-2A2;; 07/31/2001, mRNA  
sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BI566120  
BI566120.1 GI:15455312  
EST.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 542)

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B.,  
Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Friese, E.,  
George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G.,  
Misra, S., Mungall, C. J., Munoo, J., Pacleb, J., Patagas, V., Park, S.,  
Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and  
Rubin, G. M.

BDGP/HMI RH Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)

hit genomic AB003422; arm:X [1512959,1813697]

estimated-cyto:1F4-2C7: 07/31/2001

Plate: RH.352 row: D column: 10

High quality sequence stop: 471.

FEATURES  
source

1..542

Location/Qualifiers

/organism="Drosophila melanogaster"

/mol\_type="mRNA"

/db\_xref="taxon:7227"

/clone="RH35246"

/sex="male and female"

/dev\_stage="Adult"

/lab\_host="DH5-alpha Tona"

/clone\_lib="RH Drosophila melanogaster normalized Head

pF1c-1"

/note="Organ: head; Vector: pF1c1; Site 1: XhoI; Site 2:

BamHI; Library was kindly generated by Piero Carninci at

the RIKEN. The library was normalized and excised using

Cre recombinase. Plasmid cDNA library."

ORIGIN

Query Match 84.8%; Score 17.8; DB 12; Length 542;

Best Local Similarity 90.5%; Pred. No. 8e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21

|||||

49 CTGGCGTATCTGCGAGTCTG 29

RESULT 9  
BZ293209/c

LOCUS BZ293209 594 bp DNA linear GSS 31-OCT-2002

DEFINITION CG0209.f1 Candida glabrata Random Genomic Library Candida glabrata

genomic clone CG0209, genomic survey sequence.

ACCESSION BZ293209

VERSION BZ293209.1 GI:24433250

KEYWORDS GSS.

SOURCE Candida glabrata

ORGANISM Candida glabrata

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; mitosporic Saccharomycetales; Candida.

1 (bases 1 to 594)

Wong, S., Fares, M. A., Zimmermann, W., Butler, G. and Wolfe, K. H.

Evidence from comparative genomics for a complete sexual cycle in

the 'asexual' pathogenic yeast Candida glabrata

Genome Biol. 4 (2), R10 (2003)

JOURNAL

MEDLINE  
PUBMED  
COMMENT

22508158

Contact: Wong S

Department of Genetics, Smurfit Institute

Trinity College Dublin

Dublin 2, Ireland

Tel: 353 1 6082319

Fax: 353 1 6798558

Email: [swong@tcd.ie](mailto:swong@tcd.ie)

Class: plasmid ends.

FEATURES  
source

1..594

Location/Qualifiers

/organism="Candida glabrata"

/mol\_type="genomic DNA"

/strain="CBS 138"

/db\_xref="taxon:5478"

/clone="CG0209"

/clone\_lib="Candida glabrata Random Genomic Library"

ORIGIN

Query Match 84.8%; Score 17.8; DB 28; Length 594;

Best Local Similarity 90.5%; Pred. No. 8.3e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21

|||||

24 CTGGCGTATCTGAACAGTCTG 4

RESULT 10

BY719097/c

LOCUS

DEFINITION

musculus

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BY719097 631 bp mRNA linear EST 17-DEC-2002  
BY719097 RIKEN full-length enriched, adult male olfactory brain Mus  
musculus cDNA clone 6430577p16 5', mRNA sequence.

BY719097 GI:27132214

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 631)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Oatso, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojoberi, T., Baldarelli, R., Hill, D. P., Bult, C.,

Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuoka, H.,

Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusica, V.,

Chochia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,

Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,

Gariboldi, M., Giesi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,

Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,

Maltais, L., Marchionni, L., Mckenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kikukawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,

Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

22354683

PUBMED  
COMMENT

12466851

Contact: Yoshihide Hayaishizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayaishida, K., Hirozane, T., Hori, F., Imotani, K., Ighii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayaishizaki, Y.  
Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to Prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES  
source

1. 631  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="6430577P16"  
/sex="male"  
/tissue\_type="olfactory brain"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, adult male olfactory brain"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGCAACTCGAGTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATTCTCGAGTTAATAATATCCCTCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a Lambda modified pBluescript KS(+) after bulk excision from a Lambda FLC 1."

## ORIGIN

Query Match 84.8%; Score 17.8; DB 13; Length 631;  
Best Local Similarity 90.5%; Pred. No. 8.5e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CTGGCGTATCTGAAGAGTCTG 21  
|||||

Db

71 CTGGCGTATCTGAAGAGTGTG 51

## RESULT 11

AQ787069

LOCUS

DEFINITION

AQ787069 637 bp DNA linear GSS 03-AUG-1999  
HS 3156\_B2\_H06\_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3156 Col=12 Row=P, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

AQ787069.1 GI:5694693

GSS.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 637)

Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

99380589

10449764

Contact: Mahairas GS, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: <http://www.htcsc.washington.edu>

Plate: 3156 row: P column: 12

Seq primer: T7

Class: BAC ends

High quality sequence stop: 637.

Location/Qualifiers

1. 637

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="Plate=3156 Col=12 Row=P"

/sex="male"

/clone\_lib="CIT Approved Human Genomic Sperm Library D"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

E-Coli DH10B"

**AUTHORS** Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

**TITLE** Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
**MEDLINE** 99380589  
**PUBMED** 10449764

**COMMENT** Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887

Email: jwallace@u.washington.edu  
 Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.htsc.washington.edu

Plate: 3117 row: J column: 3

Seq primer: T7

Class: BAC ends

High quality sequence stop: 747.

Location/Qualifiers

1..747

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="plate=3117 Col=3 Row=J"

/sex="male"

/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in  
 E-Coli DH10B"

#### ORIGIN

Query Match 84.8%; Score 17.8; DB 28; Length 747;  
 Best Local Similarity 90.5%; Pred. No. 9.1e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21

|||||  
 419 CTGGCGTAGCTGAGAGTCTG 439

#### RESULT 13

BG030798  
**LOCUS** BG030798 817 bp mRNA linear EST 24-JAN-2001  
**DEFINITION** 602297790F1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:4392142 5',  
 mRNA sequence.

**ACCESSION** BG030798  
**VERSION** BG030798.1 GI:12420264

**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 817)

NIH-MGC http://mgc.nci.nih.gov/.

**AUTHORS** National Institutes of Health, Mammalian Gene Collection (MGC)

**TITLE** Unpublished (1999)

**JOURNAL** Unpublished (1999)

**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10084 row: c column: 23

High quality sequence stop: 725.

Location/Qualifiers

1..817

/organism="Homo sapiens"

/mol\_type="mRNA"

#### FEATURES

source

Query Match 84.8%; Score 17.8; DB 13; Length 846;  
 Best Local Similarity 90.5%; Pred. No. 9.5e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21

|||||

/db\_xref="taxon:9606"  
 /clone="IMAGE:4392142"  
 /tissue\_type="mammary adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_87"  
 /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.383 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

#### ORIGIN

Query Match 84.8%; Score 17.8; DB 10; Length 817;  
 Best Local Similarity 90.5%; Pred. No. 9.4e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21

|||||  
 765 CTGGCGTATCTGAAGAGCCTG 785

#### RESULT 14

BX730151

**LOCUS** BX730151

**DEFINITION** BX730151 XGC-tadpole Silurana tropicalis cDNA clone TTPA028h09 5',  
 mRNA sequence.

**ACCESSION** BX730151

**VERSION** BX730151.1

**KEYWORDS** EST.

**SOURCE** Silurana tropicalis

**ORGANISM** Silurana tropicalis (western clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 Xenopodinae; Silurana.

1 (bases 1 to 846)

Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.

Sanger Xenopus tropicalis EST project 2001 (11\_2003)

Unpublished (2003)

**JOURNAL** JOURNAL

**COMMENT** Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS SEQUENCE ID: TTPA028h09.plkSP6

Sequencing primer: SP6

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Nigel Garrett.

cDNA was oligo dT primed from sug of poly A+ RNA from tadpole

embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcoRI; Site 2: NotI

Host: Escherichia coli DH10B.

**FEATURES**

source

1..846

/organism="Silurana tropicalis"

/mol\_type="mRNA"

/db\_xref="taxon:8364"

/clone="TTPA028h09"

/dev\_stage="tadpole (stage 35-40)"

/lab\_host="E. coli DH10B"

/clone\_lib="XGC-tadpole"

/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
 was oligo dT primed from sug of poly A+ RNA from tadpole  
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107  
 with EcoRI at the 5' end and NotI at the 3' end"

#### ORIGIN

Db 197 CTGGCGTATCTGCAGCGTCTG 217

RESULT 15  
 BX709246  
 LOCUS  
 DEFINITION BX709246 XGC-tadpole Silurana tropicalis cDNA clone TTPA009011 5',  
 mRNA sequence.  
 ACCESSION BX709246  
 VERSION  
 KEYWORDS  
 SOURCE BX709246.1 GI:38371453  
 EST.  
 ORGANISM  
 Silurana tropicalis (western clawed frog)  
 Silurana tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Silurana.  
 REFERENCE 1 (bases 1 to 894)  
 AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.  
 TITLE Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Croning MDR  
 Sanger Institute  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: trop@sanger.ac.uk  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS\_SEQUENCE ID: TTPA009011.plkSP6  
 Sequencing primer: SP6  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Nigel Garrett.  
 cDNA was oligo dT primed from Sug of poly A+ RNA from tadpole  
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with  
 EcoRI at the 5' end and NotI at the 3' end.  
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI  
 Host: Escherichia coli DH10B.  
 FEATURES  
 source  
 1..894  
 Location/Qualifiers  
 /organism="Silurana tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone="TTPA009011"  
 /dev\_stage="tadpole (stage 35-40)"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="XGC-tadpole"  
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
 was oligo dT primed from Sug of poly A+ RNA from tadpole  
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107  
 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN  
 Query Match 84.8%; Score 17.8; DB 13; Length 894;  
 Best Local Similarity 90.5%; Pred No. 9.7e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGCAGCGTCTG 21  
 |||||  
 Db 41 CTGGCGTATCTGCAGCGTCTG 61

Search completed: April 10, 2004, 19:48:54  
 Job time : 75.8029 secs

QY 781<sup>9</sup> GCTGAGTGGGCTGGCCCTGACCAAGGTGACAGGTGACCTCTTCTACACCCATGTGAC 840  
 Db 781 GCTGAGTGGGCTGGCCCTGACCAAGGTGACAGGTGACCTCTTCTACACCCATGTGAC 840  
 QY 841 CACCATGGGCTAGAGGCTCAGCCAGAGGCCCCAGGCTGGAGGAGGTTCGATGCTTT 900  
 Db 841 CACCATGGGCTAGAGGCTCAGCCAGAGGCCCCAGGCTGGAGGAGGTTCGATGCTTT 900  
 QY 901 CAGTCAACCCAGGATGTTGGGGCACTTCAGAAAACTTCTGAGAGGTCTGTCCCTTT 960  
 Db 901 CAGTCAACCCAGGATGTTGGGGCACTTCAGAAAACTTCTGAGAGGTCTGTCCCTTT 960  
 QY 961 AGCAAGAGGCTCTGCTCCGCCCACTTCTGAGGAGGCTCTCAGCATGGGCAAGACCA 1020  
 Db 961 AGCAAGAGGCTCTGCTCCGCCCACTTCTGAGGAGGCTCTCAGCATGGGCAAGACCA 1020  
 QY 1021 GCTGGGATCATCTTCTACATGGCTGCTGTGAACAGATGCTGGAGTACCTTGTGCTG 1080  
 Db 1021 GCTGGGATCATCTTCTACATGGCTGCTGTGAACAGATGCTGGAGTACCTTGTGCTG 1080  
 QY 1081 TGGCCAGGAGCATGAGCAAAATGAACAGCAACAAAGGTGCGAGAGCAGTTGGGTTCTA 1140  
 Db 1081 TGGCCAGGAGCATGAGCAAAATGAACAGCAACAAAGGTGCGAGAGCAGTTGGGTTCTA 1140  
 QY 1141 CTCCTCGCTCTTGGGGCCATGCAAGCTGTGTGCTCTTCTACCTGCCCCCTCATTTGGCTA 1200  
 Db 1141 CTCCTCGCTCTTGGGGCCATGCAAGCTGTGTGCTCTTCTACCTGCCCCCTCATTTGGCTA 1200  
 QY 1201 CACTCATGCTGCTGAGGCTGCAAGCTGCTGAGGCGCCCACTCAGGCGACTGTCTCTCG 1260  
 Db 1201 CACTCATGCTGCTGAGGCTGCAAGCTGCTGAGGCGCCCACTCAGGCGACTGTCTCTCG 1260  
 QY 1261 AGATGCAAGGAGCGGGGTGCTACCAATCCATCAGACCAAGCTACTGCAAGATCCAAA 1320  
 Db 1261 AGATGCAAGGAGCGGGGTGCTACCAATCCATCAGACCAAGCTACTGCAAGATCCAAA 1320  
 QY 1321 GCTCAGCAATGCAATCAGTGCCTTCACTGAGCAACCTGCTGCTGCTGCTGCTGCTG 1380  
 Db 1321 GCTCAGCAATGCAATCAGTGCCTTCACTGAGCAACCTGCTGCTGCTGCTGCTGCTG 1380  
 QY 1381 CACTGCTCATCAACCACTTACACCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 Db 1381 CACTGCTCATCAACCACTTACACCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 QY 1441 TCGAGGTTTCTTCACTCAGGCTGCTGAGTCTCTATGCTGAGTGTTCCTCAATCCAA 1500  
 Db 1441 TCGAGGTTTCTTCACTCAGGCTGCTGAGTCTCTATGCTGAGTGTTCCTCAATCCAA 1500  
 QY 1501 CTTTGGGAGCTGACAGGCTGCTGCTCATCAGTGTGCTGCTGCTGCTGCTGCTGCTG 1560  
 Db 1501 CTTTGGGAGCTGACAGGCTGCTGCTCATCAGTGTGCTGCTGCTGCTGCTGCTGCTG 1560  
 QY 1561 GCCACTTTTCTGAGGAGTGGGAGCTTCAAGAGGAGGCTTCTGGGTGATCTGGG 1620  
 Db 1561 GCCACTTTTCTGAGGAGTGGGAGCTTCAAGAGGAGGCTTCTGGGTGATCTGGG 1620  
 QY 1621 CCTCTGCTATCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
 Db 1621 CCTCTGCTATCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
 QY 1681 CCGCTCCAGGAGTGGGAGCTTCAAGAGGAGGCTTCAAGAGGAGGCTTCAAGAGGAG 1740  
 Db 1681 CCGCTCCAGGAGTGGGAGCTTCAAGAGGAGGCTTCAAGAGGAGGCTTCAAGAGGAG 1740  
 QY 1741 TGAAGTACCCGATAGACTTCTCAGACCAAGGAGCTTCAAGAGGAGGCTTCAAGAGG 1800  
 Db 1741 TGAAGTACCCGATAGACTTCTCAGACCAAGGAGCTTCAAGAGGAGGCTTCAAGAGG 1800  
 QY 1801 AGCAACCAAGAGGAGTGGGAGCTTCAAGAGGAGGCTTCAAGAGGAGGCTTCAAGAGG 1860  
 Db 1801 AGCAACCAAGAGGAGTGGGAGCTTCAAGAGGAGGCTTCAAGAGGAGGCTTCAAGAGG 1860

QY 1861 GCCCTAGATTATATAATACCAAGAGAGTCTTATTTTGTAAAGCTGCAAAAAGGAGGA 1920  
 Db 1861 GCCCTAGATTATATAATACCAAGAGAGTCTTATTTTGTAAAGCTGCAAAAAGGAGGA 1920  
 QY 1921 AAAAAAACCTTCAAAAACGCCCCCTTAAGTCAACGCTCAATGACTGAAGACAGTCCCTAT 1980  
 Db 1921 AAAAAAACCTTCAAAAACGCCCCCTTAAGTCAACGCTCAATGACTGAAGACAGTCCCTAT 1980  
 QY 1981 CCTAGAGGGGTTGAGGCTTCTTCTCTGCTGGGTTGGAGGAGACAGGCTGCTTATCT 2040  
 Db 1981 CCTAGAGGGGTTGAGGCTTCTTCTCTGCTGGGTTGGAGGAGACAGGCTGCTTATCT 2040  
 QY 2041 CCTTCTAGGGGTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
 Db 2041 CCTTCTAGGGGTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
 QY 2101 GGTCCCATGTGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
 Db 2101 GGTCCCATGTGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
 QY 2161 AAAAAACACGCTCTCTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220  
 Db 2161 AAAAAACACGCTCTCTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220  
 QY 2221 AGGGGTTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT 2280  
 Db 2221 AGGGGTTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT 2280  
 QY 2281 GCTTCAAGTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2326  
 Db 2281 GCTTCAAGTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2326

RESULT 3

AAV49590  
 ID AAV49590 standard; cDNA to mRNA; 2373 BP.  
 AC AAV49590;

DT 21-OCT-1998 (first entry)

XX Human liver cell clone HPI0302 cDNA #1.

Transmembrane domain; human; nutrition; cytokine; cell proliferation; differentiation; immune system; stimulator; suppressor; regulator; hematopoiesis; activin; inhibitor; chemotactic; chemokinetic; receptor; haemostatic; thrombolytic; ligand; anti-inflammatory; tumour; ds.

OS Homo sapiens.

Key Location/Qualifiers  
 CDS 134..1813  
 /tag= a  
 /product= "transmembrane domain containing protein"

XX WO9821328-A2.

XX 22-MAY-1998.

XX 07-NOV-1997; 97WO-JP004056.

XX 13-NOV-1996; 96JP-00301429.

XX (SAGA) SAGAMI CHEM RES CENTRE.  
 (PROT-) PROTEGENE INC.

XX Kato S, Sekine S, Yamaguchi T, Kobayashi M;

XX WPI; 1998-297932/26.

XX P-PSDB; AAW64554.

XX Human protein having transmembrane domain - useful for, e.g. research and nutrition.

**THIS PAGE BLANK (USPTO)**



Query Match	95.2%	Score	2214.6	DB 2	Length	2373
Best Local Similarity	99.5%	Pred.	No. 0			
Matches 2243	Conservative	0	Mismatches	9	Indels	3
Gaps	2					
QY	64	GCTGGGGCCAGCATGGCCCCC	CACGCTGCTCAACAGGCGTACCGGAGGCGCTGTGTGATGCG	123		
DB	121	GCTGGGGCCAGCATGGCCCCC	CACGCTGCTCAACAGGCGTACCGGAGGCGCTGTGTGATGCG	180		
QY	124	CTGCACGGCTGTGCTGAGAAAC	CTTCTTCTTCTGCTGTATCTCCTGGGCTGGGGCTCCCT	183		
DB	181	CTGCACGGCTGTGCTGAGAAAC	CTTCTTCTTCTGCTGTACTCTCTGGGCTGGGGCTCCCT	240		
QY	184	GTTGATCATTTCTGAAGAAC	GAGGGCTTCTATTCCACGACGTGCCACAGCTGACGAGCAGAC	243		
DB	241	GTTGATCATTTCTGAAGAAC	GAGGGCTTCTATTCCAGCAGTGTGCCACAGCTGACGAGCAGAC	300		
QY	244	CAACACCA	CCAGGATGAGCAGCGCAGGTGGCCAGGCTGTGACCAAGCAGACGAGATGT	303		
DB	301	CAACACCA	CCAGGATGAGCAGCGCAGGTGGCCAGGCTGTGACCAAGCAGACGAGATGT	360		
QY	304	CAA	CTGGGCTTACCAATTTGGTTCTTCTGTGCTCAGGGCCACA	CAACCTTGGCATGGGAT	363	
DB	361	CAA	CTGGGCTTACCAATTTGGTTCTTCTGTGCTCAGGGCCACA	CAACCTTGGCATGGGAT	420	
QY	364	CCTCATGAC	CGGCTTTGGCCCCCGACCGCTGGCTGGTTGGCAGTGCCTGTCTCACTGC	423		
DB	421	CCTCATGAC	CGGCTTTGGCCCCCGACCGCTGGCTGGTTGGCAGTGCCTGTCTCACTGC	480		
QY	424	GTCCTGCA	CCCTCATATGCGCTTGGCTCCCGGAGCGTGGAAAGCTCTGTCTCCGTTGATATT	483		
DB	481	GTCCTGCA	CCCTCATATGCGCTTGGCTCCCGGAGCGTGGAAAGCTCTGTCTCCGTTGATATT	540		
QY	484	CCTGGCGCTGTCCCTGAA	TGGCTTTGGTTGGGCATCTGCTTAAAGCTTCACTTCACTCA	CGCT	543	
DB	541	CCTGGCGCTGTCCCTGAA	TGGCTTTGGTTGGGCATCTGCTTAAAGCTTCACTTCACTCA	CGCT	600	
QY	544	GCCCAA	CATGTTTGGGAACCTGGCTCCACGTTAA	TGGCCCTCAATGATGGCTCTTACGC	603	
DB	601	GCCCAA	CATGTTTGGGAACCTGGCTCCACGTTAA	TGGCCCTCAATGATGGCTCTTACGC	660	
QY	604	CTCTTCTGCC	ATTACGTTCCAGGAATCAAGCTGATCTACGATGCGGCTGTGGCCCTTGGT	663		
DB	661	CTCTTCTGCC	ATTACGTTCCAGGAATCAAGCTGATCTACGATGCGGCTGTGGCCCTTGGT	720		
QY	664	GGT	CATCATGTTACCTGCTTGGCTTGGCTTGGCTTATCTTCTTGA	CTGCACTGCA	CCCTCAA	723
DB	721	GGT	CATCATGTTACCTGCTTGGCTTGGCTTGGCTTATCTTCTTGA	CTGCACTGCA	CCCTCAA	780
QY	724	CTGGCCCAT	CGAAGCTTTTCTGCGCCCTTGAGGAAGTCAATTACCAAGAAAGTCAAGCT	783		
DB	781	CTGGCCCAT	CGAAGCTTTTCTGCGCCCTTGAGGAAGTCAATTACCAAGAAAGTCAAGCT	840		
QY	784	GAGTGGG	CTGGCCCTTGACACCAAGGTGACAGGTGACCTCTTCTACACCATGTGACCAAC	843		
DB	841	GAGTGGG	CTGGCCCTTGACACCAAGGTGACAGGTGACCTCTTCTACACCATGTGACCAAC	900		
QY	844	CATGGG	CCAGAGGCTCAGCCAGAAAGGCCCCAGCCTGGAGGACGGTTTGGATGCCCTTCAAT	903		
DB	901	CATGGG	CCAGAGGCTCAGCCAGAAAGGCCCCAGCCTGGAGGACGGTTTGGATGCCCTTCAAT	960		
QY	904	GTCACCC	CAGGATGTTTGGGGGCACTCAGAAAAACCTTCTCTGAGAGGTCTGTGCCCTTACG	963		

**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:11:52 ; Search time 80.7618 Seconds  
(without alignments)  
10733.545 Million cell updates/sec

Title: US-09-743-825-10  
Perfect score: 20  
Sequence: 1 gaccgatagactcttcaga 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	20	100.0	321	11	G22380	G22380 human STS W
C 2	20	100.0	457	6	AX407361	AX407361 Sequence
C 3	20	100.0	2093	6	AX880371	AX880371 Sequence
C 4	20	100.0	2093	6	BD158342	BD158342 Primer fo
C 5	20	100.0	2093	9	AK027732	AK027732 Homo sapi
C 6	20	100.0	2233	6	AX768966	AX768966 Sequence
C 7	20	100.0	2326	9	AF045584	AF045584 Homo sapi
C 8	20	100.0	2373	6	BD063631	BD063631 Human pro
C 9	20	100.0	2375	9	BC001639	BC001639 Homo sapi
C 10	20	100.0	2525	9	AB103033	AB103033 Homo sapi
C 11	20	100.0	136678	2	AP001332	AP001332 Homo sapi
C 12	20	100.0	143899	2	AC013606	AC013606 Homo sapi
C 13	20	100.0	143899	6	AX411543	AX411543 Sequence
C 14	20	100.0	165434	2	AC087505	AC087505 Homo sapi
C 15	20	100.0	178000	9	AP002893	AP002893 Homo sapi
C 16	20	100.0	187566	2	AC015685	AC015685 Homo sapi
C 17	18.4	92.0	248825	2	AC132054	AC132054 Rattus no
C 18	18.4	92.0	260186	2	AC111920	AC111920 Rattus no
C 19	17.4	87.0	181595	2	AC136121	AC136121 Rattus no
C 20	17.4	87.0	222814	10	AC134591	AC134591 Mus muscu
C 21	16.8	84.0	2337	9	AB070000	AB070000 Macaca fa
C 22	16.8	84.0	38826	9	AC005356	AC005356 Homo sapi
C 23	16.8	84.0	59866	2	AC100572	AC100572 Mus muscu
C 24	16.8	84.0	110000	2	AC095743	Continuation (2 of
C 25	16.8	84.0	156075	10	AL844157	AL844157 Mouse DNA
C 26	16.8	84.0	159818	2	AC141050	AC141050 Rattus no
C 27	16.8	84.0	163511	9	AL442203	AL442203 Human DNA
C 28	16.8	84.0	168194	2	AC141475	AC141475 Mus muscu
C 29	16.8	84.0	172071	9	AC012676	AC012676 Homo sapi
C 30	16.8	84.0	175801	10	AC132464	AC132464 Mus muscu
C 31	16.8	84.0	180745	9	AL359183	AL359183 Human DNA
C 32	16.8	84.0	182679	9	AL365207	AL365207 Human DNA
C 33	16.8	84.0	183457	2	AC128106	AC128106 Rattus no
C 34	16.8	84.0	195582	2	AC105486	AC105486 Rattus no
C 35	16.8	84.0	196545	9	AC113390	AC113390 Rattus no
C 36	16.8	84.0	212712	2	AC135641	AC135641 Mus muscu
C 37	16.8	84.0	213331	10	AC098745	AC098745 Mus muscu
C 38	16.8	84.0	214779	2	AC112769	AC112769 Rattus no
C 39	16.8	84.0	232067	2	AC121286	AC121286 Mus muscu
C 40	16.8	84.0	250620	2	AC137946	AC137946 Mus muscu
C 41	16.8	84.0	252675	2	AC116184	AC116184 Rattus no
C 42	16.8	84.0	253140	2	AC094055	AC094055 Rattus no
C 43	16.8	84.0	254575	2	AC127999	AC127999 Rattus no
C 44	16.8	84.0	255436	2	AC125922	AC125922 Rattus no
C 45	16.8	84.0	265985	2	AC136674	AC136674 Rattus no

# ALIGNMENTS

RESULT 1  
G22380/c  
LOCUS  
DEFINITION human STS WI-17004, sequence tagged site.  
ACCESSION G22380  
VERSION G22380.1 GI:1342706  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 321)  
AUTHORS Hudson.T.  
TITLE Whitehead Institute/MIT Center for Genome Research; Physically  
Mapped STS

**THIS PAGE BLANK (USPTO)**

Thu Apr 15 09:04:00 2004

JOURNAL  
COMMENT

Unpublished (1995)

Contact: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu

Primer A: TTTATTAATCTACGCCATGGC  
Primer B: CTGGATGACAGCATCAAG  
STS size: 101  
PCR Profile:

Presoak:  
Denaturation:  
Annealing: 56 degrees C  
Polymerization:  
PCR Cycles: 35  
Thermal Cycler:  
Protocol:  
Template: 10 ng  
Primer: each 5 pM  
dNTPs: each 4 nM  
Taq Polymerase: 0.025 units/ul  
Total Vol: 20 ul

Buffer:

MgCl2: 1.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 9.3

Derived from dbEST (genbank accession R00504).  
Location/Qualifiers  
1..321

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="291.1 CR from top of Chr11 linkage group"  
35..135  
35..135  
Complement (116..135)

Query Match

Best Local Similarity 100.0%; Score 20; DB 11; Length 321;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
DB 164 GACCGCATAGACTTCTCAGA 145

RESULT 2  
AX407361/c

LOCUS AX407361  
DEFINITION Sequence 8 from Patent WO0229103.  
ACCESSION AX407361  
VERSION AX407361.1 GI:21440066  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Alvares, C., Horne, D., Peres-da-Silva, S.  
Gene expression

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
STS  
primer\_bind  
primer\_bind  
ORIGIN  
Query Match  
Best Local Similarity 100.0%; Score 20; DB 11; Length 321;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

us-09-743-825-10.rge

Page 2

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="EMBL/GenBank Accession No. AA004521"

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 457;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
DB 144 GACCGCATAGACTTCTCAGA 125

RESULT 3  
AX880371

LOCUS AX880371  
DEFINITION Sequence 15276 from Patent EP1074617.  
ACCESSION AX880371  
VERSION AX880371.1 GI:40035107  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Ota, T., Sugai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
Primers for synthesizing full-length cDNA and their use  
Patent: EP 1074617-A 15276 07-FEB-2001;  
Research Association for Biotechnology (JP)  
Location/Qualifiers  
1..2093

REFERENCE  
AUTHORS

TITLE  
JOURNAL

FEATURES  
source

CDS

/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
77..1534  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAE90921.1"  
/db\_xref="GI:40035108"  
/translation="MLNLGPTIGSFVLSATPLGLMDRFPQRPVRLVGSACFTASC  
SSATFPGLKLYDAGVAFVIMPTWSGLACILFNCTLNWPIERFAPPEEVNYYTKI  
KLSGLADHKVTGDLFTVHTMTCQRSLSDGSDAEMSFDVVGTSNLPERS  
VPLKSLCSPTFLWSLLTMGTQRIIFYNAAVNKMLEYLVTGGQEHETNEQQQKVAR  
RYCKIQKLTWASITSAFTLNLVGFGITCLNNLHQFVTFVLTIVRGPVHSAGSL  
YAAVFPNSHFTLTLQSLISAVFALLQQLPLFMVGLKGEFPFWNLGLLFLSLGFP  
LLPSYLYFVRLAQEYAAANGMGLKVLGSEVTA"

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 2093;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
DB 1525 GACCGCATAGACTTCTCAGA 1544

RESULT 4  
BD158342

LOCUS BD158342

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

STS

primer\_bind

primer\_bind

**THIS PAGE BLANK (USPTO)**



Thu Apr 15 09:04:00 2004

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: April 10, 2004, 11:11:52 ; Search time 88.838 Seconds  
(without alignments)  
10733.545 Million cell updates/sec

Title: US-09-743-825-7  
Perfect score: 22  
Sequence: 1 gcattgtacagtagaagcc 22

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0  
Maximum Match 100  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: gb\_vl.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_hgt\_hum.\*  
40: em\_hgt\_mus.\*  
41: .\*

us-09-743-825-7.rge

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	321	11	G22380	G22380 human STS W
2	22	100.0	457	6	AX407361	AX407361 Sequence
3	22	100.0	2093	6	AX880371	AX880371 Sequence
4	22	100.0	2093	6	BD158342	BD158342 Primer fo
5	22	100.0	2093	9	AK027732	AK027732 Homo sapi
6	22	100.0	2233	6	AX768966	AX768966 Sequence
7	22	100.0	2326	9	AF045584	AF045584 Homo sapi
8	22	100.0	2373	6	BD063631	BD063631 Homo sapi
9	22	100.0	2375	9	BC001639	BC001639 Human pro
10	22	100.0	2525	9	AB103033	AB103033 Homo sapi
11	22	100.0	136678	2	AP001332	AP001332 Homo sapi
12	22	100.0	143899	2	AC013606	AC013606 Homo sapi
13	22	100.0	143899	6	AX111543	AX111543 Sequence
14	22	100.0	165434	2	AC087505	AC087505 Homo sapi
15	22	100.0	178000	9	AF002893	AF002893 Homo sapi
16	22	100.0	187566	2	AC015685	AC015685 Homo sapi
17	19.4	88.2	601	11	G88016	G88016 S209P6569PE
18	19.4	88.2	192042	9	AC079955	AC079955 Homo sapi
19	19.4	88.2	243654	10	AL663082	AL663082 Mouse DNA
20	19	86.4	83709	2	AC144727	AC144727 Medicago
21	19	86.4	102421	2	AC146755	AC146755 Medicago
22	19	86.4	102858	2	AC146556	AC146556 Medicago
23	19	86.4	103039	8	AC139525	AC139525 Medicago
24	19	86.4	108694	2	AC135848	AC135848 Medicago
25	19	86.4	118705	8	AC126009	AC126009 Medicago
26	19	86.4	125883	8	AC124959	AC124959 Medicago
27	19	86.4	126687	2	AC131240	AC131240 Medicago
28	19	86.4	140418	2	AC137822	AC137822 Medicago
29	19	86.4	248040	2	AC137521	AC137521 Medicago
30	18.8	85.5	579	6	AX874028	AX874028 Sequence
31	18.8	85.5	579	6	BD154090	BD154090 Primer fo
32	18.8	85.5	171673	5	AL928790	AL928790 Zebrafish
33	18.8	85.5	181862	5	BX255897	BX255897 Zebrafish
34	18.4	83.6	47069	9	AC021101	AC021101 Homo sapi
35	18.4	83.6	156220	2	AC138738	AC138738 Mus muscu
36	18.4	83.6	192015	2	AC119223	AC119223 Mus muscu
37	18.4	83.6	199670	2	AC144802	AC144802 Mus muscu
38	18	81.8	230127	10	AL691481	AL691481 Mouse DNA
39	18	81.8	348498	1	BX571872	BX571872 Photorhab
40	18	81.8	349980	6	AX770910	AX770910 Sequence
41	17.8	80.9	2476	8	KLKLEU2G	KLKLEU2G
42	17.8	80.9	43051	10	BX649238	BX649238 Mouse DNA
43	17.8	80.9	50374	2	AC015170	AC015170 Drosophil
44	17.8	80.9	63839	9	AL139394	AL139394 Human DNA
45	17.8	80.9	97065	2	AP000768	AP000768 Homo sapi

ALIGNMENTS

RESULT 1  
G22380  
LOCUS  
DEFINITION human STS WI-17004, sequence tagged site.  
ACCESSION G22380  
VERSION G22380.1  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE capiens (human)  
G22380 human STS WI-17004, sequence tagged site.  
G22380  
G22380.1 GI:1342706  
STS; STS sequence; primer; sequence tagged site.  
capiens (human)

**THIS PAGE BLANK (USPTO)**

Thu Apr 15 09:04:00 2004

JOURNAL  
COMMENT

Unpublished (1995)

Contact: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu

Primer A: TTATTAATCTACGGCCATGCG  
Primer B: CTGATGACAGGCAATCAAG  
STS size: 101  
PCR Profile:

Presoak:  
Denaturation:  
Annealing: 56 degrees C  
Polymerization:  
PCR Cycles: 35  
Thermal Cycler:

Protocol:  
Template: 10 ng  
Primer: each 5 pm  
dNTPs: each 4 mM  
Tag Polymerase: 0.025 units/ul  
Total Vol: 20 ul

Buffer:

MgCl2: 1.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 9.3

Derived from dbEST (genbank accession R00504).  
Location/Qualifiers  
1..321

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="291.1 CR from top of Chr11 linkage group"  
35..135  
35..156  
complement(116..135)

STIS

primer\_bind  
primer\_bind

Query Match  
Best Local Similarity 100.0%; Score 22; DB 11; Length 321;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAGCC 22

Db 64 GCATGTTACAGGTAGAAAGCC 85

RESULT 2  
AX407361

LOCUS  
DEFINITION  
Sequence 8 from Patent W00229103. 457 bp  
Accession AX407361 DNA linear PAT 14-JUN-2002

VERSION  
AX407361.1 GI:21440066

KEYWORDS  
SOURCE

ORGANISM  
Homo sapiens (human)

REFERENCE  
AUTHORS  
Alvares, C., Horne, D., Peres-da-Silva

TITLE  
Gene expression

JOURNAL

us-09-743-825-7.rge

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="EMBL/GenBank Accession No. AA004521"

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 22; DB 6; Length 457;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAGCC 22

Db 44 GCATGTTACAGGTAGAAAGCC 65

RESULT 3  
AX880371/c

LOCUS  
DEFINITION  
Sequence 15276 from Patent EP1074617. 2093 bp  
Accession AX880371 DNA linear PAT 17-DEC-2003

VERSION  
AX880371.1 GI:40035107

KEYWORDS  
SOURCE

ORGANISM  
Homo sapiens (human)

REFERENCE  
AUTHORS  
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.

TITLE  
JOURNAL  
Primers for synthesizing full-length cDNA and their use

FEATURES  
source  
Location/Qualifiers  
1..2093

/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
77..1534

/note="unnamed protein product"  
/protein\_id="CAE90921.1"  
/translation="MLNLGFTIGSFVLSATLPLGILMDRPPVRLVGSACFTASC  
SSAITPGIKLIYDAGVAVVIMFTWSGLACFLNCTLNWPIEAFPAPEVNTYTKI  
KLSGLADHKVTGDLFVHTVTTMGRLSKAPSDGSDAPMSQDVRGTSENLPERS  
TVGFTSVFGMQLLCTCTPLIGYIMDWIRKDCVADPTQGTGVDHARDGVAATKSRP  
YAAVPSNHFPTGLTQSLISAVFALLQPLFMAMVGLKGPFWNLGLLFLSLG  
LLPSLYFYTRARLQOEYAANGMGLKVLGSGSEVTA"

Query Match  
Best Local Similarity 100.0%; Score 22; DB 6; Length 2093;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAGCC 22

Db 1625 GCATGTTACAGGTAGAAAGCC 1604

RESULT 4  
ED15834

**THIS PAGE BLANK (USPTO)**



THIS PAGE BLANK (USPTO)

XX	PS	Claim 4; Page 166-169; 205pp; English.	
XX	CC	AAV49550-V49599 are cDNA sequences which encode human proteins containing a transmembrane domain. These proteins can be used for, e.g. research and nutrition, and may have cytokine and cell proliferation/differentiation, immune stimulating/suppressing, haematopoiesis regulating, tissue growth, activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic, receptor/ligand, anti-inflammatory or tumour inhibition activity	
XX	SQ	Sequence 2373 BP; 469 A; 716 C; 645 G; 543 T; 0 U; 0 Other;	
		Query Match 95.2%; Score 2214.6; DB 2; Length 2373;	
		Best Local Similarity 99.5%; Pred. No. 0;	
		Matches 2243; Conservative 0; Mismatches 9; Indels 3; Gaps 2;	
QY	64	GCTCGGGGCGCAGCATGGCCCGCCACGCTGCAACAGGCGGTACCGAGGCGCTGGTGGATGGC	123
QY	121	GCTCGGGGCGCAGCATGGCCCGCCACGCTGCAACAGGCGGTACCGAGGCGCTGGTGGATGGC	180
QY	124	CTGACGGGTGCTGGAGAACCTCTTCTCTGCTGTACTCTCTGGGCTGGGGTCCCT	183
DB	181	CTGACGGGTGCTGGAGAACCTCTTCTCTGCTGTACTCTCTGGGCTGGGGTCCCT	240
QY	184	GTTGATCATCTGAGAACGAGGGCTTCTATTCCAGCAGCTGCCAGCTGAGGAGCAGCAC	243
DB	241	GTTGATCATCTGAGAACGAGGGCTTCTATTCCAGCAGCTGCCAGCTGAGGAGCAGCAC	300
QY	244	CAACACACCCAGGATGAGCAGGCGAGGTGGCGAGCTGTGACCCAGCAGGACGAGATGCT	303
DB	301	CAACACACCCAGGATGAGCAGGCGAGGTGGCGAGCTGTGACCCAGCAGGACGAGATGCT	360
QY	304	CAACCTGGGCTTCCATTTGGTTCCTTCTGCTGTGCTGCTGAGGCGCACCCCTCCGCTGAGGAT	363
DB	361	CAACCTGGGCTTCCATTTGGTTCCTTCTGCTGTGCTGCTGAGGCGCACCCCTCCGCTGAGGAT	420
QY	364	CCTCATGAGCCGCTTTGGCCCGCCGACCGTGGCGCTGGTGGCAGTCCCTTCACTGCG	423
DB	421	CCTCATGAGCCGCTTTGGCCCGCCGACCGTGGCGCTGGTGGCAGTCCCTTCACTGCG	480
QY	424	GTCTCTGACCCCTCATGGCCCTTGCCCTCCCGGAGCTGGAAGCTCTCTCCGTTGATATT	483
DB	481	GTCTCTGACCCCTCATGGCCCTTGCCCTCCCGGAGCTGGAAGCTCTCTCCGTTGATATT	540
QY	484	CTGGGCGCTGCTCCGATGGCTTTGGTGGCATCTGCTGCTAAAGCTTCACTACGCT	543
DB	541	CTGGGCGCTGCTCCGATGGCTTTGGTGGCATCTGCTGCTAAAGCTTCACTACGCT	600
QY	544	GCCCAACATGTTGGGAACTGGCGCTCCAGTTAATGGCCCTCATGATGGCTCTTACGC	603
DB	601	GCCCAACATGTTGGGAACTGGCGCTCCAGTTAATGGCCCTCATGATGGCTCTTACGC	660
QY	604	CTCTCTGCAATACGTTCCAGGAATCAAGCTGATCTAGCATGCGGCTGGCCCTTCT	663
DB	661	CTCTCTGCAATACGTTCCAGGAATCAAGCTGATCTAGCATGCGGCTGGCCCTTCT	720
QY	664	GGTGATCATGTTCACTGGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	723
DB	721	GGTGATCATGTTCACTGGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	780
QY	724	CTGGCCCATGAGCCCTTCTGCGCCCTGAGGAGTCAATTAACGAGAGATCAAGCT	783
DB	781	CTGGCCCATGAGCCCTTCTGCGCCCTGAGGAGTCAATTAACGAGAGATCAAGCT	840
QY	784	GAGTGGGCTGGCCCTGAGAACCAAGGTGACAGGTGACCTCTCTACACCCATGTGACCCAC	843
DB	841	GAGTGGGCTGGCCCTGAGAACCAAGGTGACAGGTGACCTCTCTACACCCATGTGACCCAC	900
QY	844	CATGGGCGCAGGCTCAGCCAGAGGCGCCCGAGCTGGAGGACGGTTCGGATCCCTTCAT	903
DB	901	CATGGGCGCAGGCTCAGCCAGAGGCGCCCGAGCTGGAGGACGGTTCGGATCCCTTCAT	960
QY	904	GTCAACCCAGGATGTTCTGGGGGCACTTCAGAAAAACCTTCTGAGAGGTCTGTCCCTTACG	963

DB	961	GTCAACCCAGGATGTTGGGGCACCTCAGAAAACTTCTCTGAGAGGTCTGTCCCTTACG	1020
QY	964	CAAGAGCCTCTGCTCCCGCCACCTTTCTGTGGAGCCTCTCTCAACATGGGATGAGCCAGCT	1023
DB	1021	CAAGAGCCTCTGCTCCCGCCACCTTTCTGTGGAGCCTCTCTCAACATGGGATGAGCCAGCT	1080
QY	1024	CGGATCATCTTCTACATGGCTGTGTGAACAAAGATGCTGAGTACCTTTGTGACTGGGG	1083
DB	1081	CGGATCATCTTCTACATGGCTGTGTGAACAAAGATGCTGAGTACCTTTGTGACTGGGG	1140
QY	1084	CCAGGACATGAGACAAATGAACAGCAACAAAGGTGGCAGAGACAGTTGGTCTTACTC	1143
DB	1141	CCAGGACATGAGACAAATGAACAGCAACAAAGGTGGCAGAGACAGTTGGTCTTACTC	1200
QY	1144	CTCCGCTTTCGGGCGCATGAGCTGTGTGCTCTCTCTCACTGCGCCCTCATTTGGGTACAT	1203
DB	1201	CTCCGCTTTCGGGCGCATGAGCTGTGTGCTCTCTCTCACTGCGCCCTCATTTGGGTACAT	1260
QY	1204	CATGGAATGGCGGATCAGGACTGGTGGAGCGCCCACTCAGGCACTGCTCTCGGAGA	1263
DB	1261	CATGGAATGGCGGATCAGGACTGGTGGAGCGCCCACTCAGGCACTGCTCTCGGAGA	1320
QY	1264	TGCCAGGACCGGGTGTGTACAAATCATCAGACCAAGCTACTGCAAGATCCAAAGCT	1323
DB	1321	TGCCAGGACCGGGTGTGTACAAATCATCAGACCAAGCTACTGCAAGATCCAAAGCT	1380
QY	1324	CACCAATGCCATCAGTGCCTTTCACCTGACCAACTGCTGTGTGGGTGGGATCAC	1383
DB	1381	CACCAATGCCATCAGTGCCTTTCACCTGACCAACTGCTGTGTGGGTGGGATCAC	1440
QY	1384	CTGTCTCATGACCAACTTACACCTCCAGTGTGTGACCTTCTGCTGACACCATGTTGCG	1443
DB	1441	CTGTCTCATGACCAACTTACACCTCCAGTGTGTGACCTTCTGCTGACACCATGTTGCG	1500
QY	1444	AGGTTTCTTCCACTCAGCCTGTGGGAGTCTCTATGTGCTGAGTGTTCCTATCCAACTT	1503
DB	1501	AGGTTTCTTCCACTCAGCCTGTGGGAGTCTCTATGTGCTGAGTGTTCCTATCCAACTT	1560
QY	1504	TGGGACGCTGACAGGCTGCGCTCTCATCAGTGTGTGTGTGCTGCTTTCAGCAGCC	1563
DB	1561	TGGGACGCTGACAGGCTGCGCTCTCATCAGTGTGTGTGTGCTGCTTTCAGCAGCC	1620
QY	1564	ACTTTTCTATGCGGATGGTGGGACCCCTCAAGAGGAGCCCTTCTGGGTGAATCTGGGCT	1623
DB	1621	ACTTTTCTATGCGGATGGTGGGACCCCTCAAGAGGAGCCCTTCTGGGTGAATCTGGGCT	1680
QY	1624	CCTGCTATTTCTCACTCTCTGGGATTCCTGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCT	1683
DB	1681	CCTGCTATTTCTCACTCTCTGGGATTCCTGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCT	1740
QY	1684	GCTCAGCAGGAGTACGCGCCCAATGGGATGGGCGCACTGAAGTGTCTTGGGCTCTGA	1743
DB	1741	GCTCAGCAGGAGTACGCGCCCAATGGGATGGGCGCACTGAAGTGTCTTGGGCTCTGA	1800
QY	1744	GSTGACCGATAGACTTCTCAGACCAAGGACCTGGATGACAGGCAATCAAGGCTGAGC	1803
DB	1801	GSTGACCGATAGACTTCTCAGACCAAGGACCTGGATGACAGGCAATCAAGGCTGAGC	1860
QY	1804	AACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTGTAACTGACATAGAGCCATGGCC	1863
DB	1861	AACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTGTAACTGACATAGAGCCATGGCC	1920
QY	1864	GTAGATTTTAAATACCAAGGAGTCTTATTTTGTAAAGTCTGCAAAAGAGGAGG-AAA	1922
DB	1921	GTAGATTTTAAATACCAAGGAGTCTTATTTTGTAAAGTCTGCAAAAGAGGAGGAGGAAA	1980
QY	1923	AAAAACCTTCAAAAACGCCCCCTTAAGTCAACGCTCATTGACTGAGAGTCCCTATCC	1982
DB	1981	AAAAACCTTCAAAAACGCCCCCTTAAGTCAACGCTCATTGACTGAGAGTCCCTATCC	2040
QY	1983	TAGAGGGGTGAGCTTTCTTCTCTCTTGGGTGGAGGAGACAGGCTGCTCTTATCTCC	2042

THIS PAGE BLANK (USPTO)





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: April 10, 2004, 11:11:52 ; Search time 80.7618 Seconds  
(without alignments)  
10733.545 Million cell updates/sec

Title: US-09-743-825-10  
Perfect score: 20  
Sequence: 1 gaccgcagatcttcaga 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.ste.\*
- 12: gb.sv.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.in.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.mu.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.ste.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_mam.\*
- 37: em.htg\_vrt.\*
- 38: em\_sy.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	20	100.0	321	11	G22380	G22380 human STS W
C 2	20	100.0	457	6	AX407361	AX407361 Sequence
C 3	20	100.0	2093	6	AX880371	AX880371 Sequence
C 4	20	100.0	2093	6	BD158342	BD158342 Primer fo
C 5	20	100.0	2093	9	AX027732	AX027732 Homo sapi
C 6	20	100.0	2233	6	AX768966	AX768966 Sequence
C 7	20	100.0	2326	9	AF045584	AF045584 Homo sapi
C 8	20	100.0	2373	6	BD063631	BD063631 Human pro
C 9	20	100.0	2375	9	BC001639	BC001639 Homo sapi
C 10	20	100.0	2525	9	AB103033	AB103033 Homo sapi
C 11	20	100.0	136678	2	AP001332	AP001332 Homo sapi
C 12	20	100.0	143899	2	AC013606	AC013606 Homo sapi
C 13	20	100.0	143899	6	AX411543	AX411543 Sequence
C 14	20	100.0	165434	2	AC087505	AC087505 Homo sapi
C 15	20	100.0	178000	9	AP002893	AP002893 Homo sapi
C 16	20	100.0	187565	2	AC015685	AC015685 Homo sapi
C 17	18.4	92.0	248825	2	AC132054	AC132054 Rattus no
C 18	18.4	92.0	260186	2	AC111920	AC111920 Rattus no
C 19	17.4	87.0	181595	2	AC136121	AC136121 Rattus no
C 20	17.4	87.0	222814	10	AC134591	AC134591 Mus muscu
C 21	16.8	84.0	2337	9	AB070000	AB070000 Macaca fa
C 22	16.8	84.0	38826	9	AC005356	AC005356 Homo sapi
C 23	16.8	84.0	59866	2	AC100572	AC100572 Mus muscu
C 24	16.8	84.0	110000	2	AC095743	Continuation (2 of
C 25	16.8	84.0	156075	10	AL844157	AL844157 Mouse DNA
C 26	16.8	84.0	159818	2	AC141050	AC141050 Rattus no
C 27	16.8	84.0	163511	9	AL442203	AL442203 Human DNA
C 28	16.8	84.0	168194	2	AC141475	AC141475 Mus muscu
C 29	16.8	84.0	172071	9	AC012676	AC012676 Homo sapi
C 30	16.8	84.0	175801	10	AC132464	AC132464 Mus muscu
C 31	16.8	84.0	180745	9	AL359183	AL359183 Human DNA
C 32	16.8	84.0	182679	9	AL365207	AL365207 Human DNA
C 33	16.8	84.0	183457	2	AC128106	AC128106 Rattus no
C 34	16.8	84.0	195582	2	AC105486	AC105486 Rattus no
C 35	16.8	84.0	196545	9	AC113390	AC113390 Homo sapi
C 36	16.8	84.0	212712	2	AC135641	AC135641 Mus muscu
C 37	16.8	84.0	213331	10	AC098745	AC098745 Mus muscu
C 38	16.8	84.0	214779	2	AC112769	AC112769 Rattus no
C 39	16.8	84.0	232067	2	AC121286	AC121286 Mus muscu
C 40	16.8	84.0	250620	2	AC137946	AC137946 Mus muscu
C 41	16.8	84.0	252675	2	AC116184	AC116184 Rattus no
C 42	16.8	84.0	253140	2	AC094055	AC094055 Rattus no
C 43	16.8	84.0	254575	2	AC127999	AC127999 Rattus no
C 44	16.8	84.0	255436	2	AC125922	AC125922 Rattus no
C 45	16.8	84.0	265985	2	AC136674	AC136674 Rattus no

ALIGNMENTS

RESULT 1  
G22380/c  
LOCUS human STS WI-17004, sequence tagged site.  
DEFINITION G22380  
ACCESSION G22380  
VERSION G22380.1 GI:1342706  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 321)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE Hudson, T.  
Whitehead Institute/MIT Center for Genome Research; Physically  
Mapped STGS

321 bp DNA linear STS 31-MAY-1996

**THIS PAGE BLANK (USPTO)**

JOURNAL  
COMMENT

Unpublished (1995)

Contact: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
Whitehead Institute for Biomedical Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu

Primer A: TTTATAAATCTACGGCCATGGC  
Primer B: CTGGATGACAGGCAATCAAG  
STS size: 101  
PCR Profile:

Prisoak:  
Denaturation:  
Annealing: 56 degrees C  
Polymerization:  
PCR Cycles: 35  
Thermal Cycler:  
Protocol:  
Template: 10 ng  
Primer: each 5 pM  
dNTPs: each 4 mM  
Taq Polymerase: 0.025 units/ul  
Total Vol: 20 ul

## Buffer:

MgCl2: 1.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 9.3

Derived from dbEST (genbank accession R00504).

FEATURES  
source  
Location/Qualifiers  
1..321  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="291.1 CR from top of Chr11 linkage group"  
35..135  
35..135  
primer\_bind  
primer\_bind  
complement(116..135)

## ORIGIN

Query Match 100.0%; Score 20; DB 11; Length 321;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 164 GACCGCATAGACTTCTCAGA 145

## RESULT 2

AX407361/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1  
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.  
Gene expression profiles in liver cancer  
Patent: WO 0229103-A 8 11-APR-2002;  
GENE LOGIC INC (US)  
Location/Qualifiers  
1..457  
/organism="Homo sapiens"

FEATURES  
source

Location/Qualifiers  
1..457  
/organism="Homo sapiens"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 144 GACCGCATAGACTTCTCAGA 125

## RESULT 3

AX880371  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1  
Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
Primers for synthesizing full-length cDNA and their use  
Patent: EP 1074617-A 15276 07-FEB-2001;  
Research Association for Biotechnology (JP)  
Location/Qualifiers  
1..2093  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
77..1534  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAE90921.1"  
/db\_xref="GI:40035108"  
/translation="MLNLGFTIGSFVLSATLLPLGILMDRFGPRPRLVLSACFTASC  
TMAASRDVEALSPLIFLALSLNGFGICLTFTSLTPNNFGLRSLTMAIMGSYA  
SSAITFGKLIYDAGVAFVIMFTWSGLACLIPLNCTLNWPIFAFPAPPEVYTKKI  
KSLGLALDHKVTGDLFYHTVTMGQRSLQKAPSLDGSADAFMSPODVRGTSENLPERS  
VPLRKSICSPFTFLSLLTMGTOLRIIFMAAVNKMLEYLVTSQEHETNSQQKVAE  
TVGVYSSVFGAMQLLCILTCPLIGYIMDRIKDCVDAPTQSTVLGDARDGVATKSRP  
RYCKIKLTNAISAFITLNLVGFITCLINNLHLQFVTFVLTIVRGFFHLLSLLGF  
YAAVFPNHFGLTGLQLSLISAVFALLOQPLFMAMVGLKGEPTFWNLGLLFLSLGFL  
LLPSYLFYRRLQOEYAANGMPLKVLGSEVTA"

## FEATURES

source

## CDS

Location/Qualifiers  
77..1534  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAE90921.1"  
/db\_xref="GI:40035108"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2093;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1525 GACCGCATAGACTTCTCAGA 1544

## RESULT 4

BD158342  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 2093)

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:11:52 ; Search time 88.838 seconds  
(without alignments)  
10733.545 Million cell updates/sec

Title: US-09-743-825-7

Perfect score: 22

Sequence: 1 gcattgtacaggtagaaagcc 22

Scoring table:

IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_ats.\*

12: gb\_ay.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_ats.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	321	11	G22380	G22380 human STS W
2	22	100.0	457	6	AX407361	AX407361 Sequence
3	22	100.0	2093	6	AX880371	AX880371 Sequence
4	22	100.0	2093	6	BD158342	BD158342 Primer fo
5	22	100.0	2093	9	AK027732	AK027732 Homo sapi
6	22	100.0	2233	6	AX768966	AX768966 Sequence
7	22	100.0	2326	9	AF045584	AF045584 Homo sapi
8	22	100.0	2373	6	BD063631	BD063631 Human pro
9	22	100.0	2375	6	BC001639	BC001639 Homo sapi
10	22	100.0	2525	9	AB103033	AB103033 Homo sapi
11	22	100.0	136678	2	AP001332	AP001332 Homo sapi
12	22	100.0	143899	6	AX411543	AX411543 Sequence
13	22	100.0	143899	6	AX411543	AX411543 Sequence
14	22	100.0	165434	2	AC087505	AC087505 Homo sapi
15	22	100.0	178000	9	AP002893	AP002893 Homo sapi
16	22	100.0	187566	2	AC015685	AC015685 Homo sapi
17	19.4	88.2	601	11	G88016	G88016 S208P6569FE
18	19.4	88.2	192042	9	AC079955	AC079955 Homo sapi
19	19.4	88.2	243654	10	AL663082	AL663082 Mouse DNA
20	19	86.4	83709	2	AC144727	AC144727 Medicago
21	19	86.4	102421	2	AC146755	AC146755 Medicago
22	19	86.4	102858	2	AC144656	AC144656 Medicago
23	19	86.4	103039	8	AC139525	AC139525 Medicago
24	19	86.4	108694	2	AC135848	AC135848 Medicago
25	19	86.4	118705	2	AC136009	AC136009 Medicago
26	19	86.4	125883	8	AC124959	AC124959 Medicago
27	19	86.4	126687	2	AC131240	AC131240 Medicago
28	19	86.4	140418	2	AC137822	AC137822 Medicago
29	19	86.4	248040	2	AC137521	AC137521 Medicago
30	18.8	85.5	579	6	AX874028	AX874028 Sequence
31	18.8	85.5	579	6	BD154090	BD154090 Primer fo
32	18.8	85.5	171673	5	BL928790	BL928790 Zebrafish
33	18.8	85.5	181862	5	BX255897	BX255897 Zebrafish
34	18.4	83.6	47069	9	AC021101	AC021101 Homo sapi
35	18.4	83.6	156220	2	AC138738	AC138738 Mus muscu
36	18.4	83.6	192015	2	AC119223	AC119223 Mus muscu
37	18.4	83.6	199670	2	AC144802	AC144802 Mus muscu
38	18	81.8	230127	10	AL691481	AL691481 Mouse DNA
39	18	81.8	348498	1	BX571872	BX571872 Phototab
40	18	81.8	349980	6	AX770910	AX770910 Sequence
41	17.8	80.9	2476	8	KLKLEU2G	X65545 K.lactis KL
42	17.8	80.9	43051	10	BX649238	BX649238 Mouse DNA
43	17.8	80.9	50374	2	AC015170	AC015170 Drosophil
44	17.8	80.9	63839	9	AL139394	AL139394 Human DNA
45	17.8	80.9	97085	2	AP000768	AP000768 Homo sapi

# ALIGNMENTS

RESULT 1  
G22380  
LOCUS  
DEFINITION human STS WI-17004, sequence tagged site.  
ACCESSION G22380  
VERSION G22380.1 GI:1342706  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 321)  
Hudson, T.  
Whitehead Institute/MIT Center for Genome Research; Physically  
Mapped STSs

Pred. No. is the number of results predicted by chance to have a

**THIS PAGE BLANK (USPTO)**



JOURNAL  
COMMENT

## Unpublished (1995)

Contact: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
Whitehead Institute for Biomedical Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu

Primer A: TTTGATTAATCTACGGCCATGGC  
Primer B: CTGATCAGGCAATCAAG  
STS size: 101  
PCR Profile:

Presoak:  
Denaturation: 56 degrees C  
Annealing: 56 degrees C  
Polymerization:  
PCR Cycles: 35  
Thermal Cycler:

## Protocol:

Template: 10 ng  
Primer: each 5 pM  
dNTPs: each 4 nM  
Taq Polymerase: 0.025 unite/ul  
Total Vol: 20 ul

## Buffer:

MgCl2: 1.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 9.3

Derived from dbEST (genbank accession R00504).

FEATURES  
source

1..321  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="291.1 cR from top of Chr11 linkage group"  
35..135  
35..56  
primer\_bind  
primer\_bind  
complement(116..135)

## ORIGIN

Query Match 100.0%; Score 22; DB 11; Length 321;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
DB 64 GCATGTTACAGGTAGAAAAGCC 85

RESULT 2  
AX407361

LOCUS AX407361 457 bp DNA linear PAT 14-JUN-2002  
DEFINITION Sequence 8 from Patent WO0229103.  
ACCESSION AX407361  
VERSION AX407361.1 GI:21440066

KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.  
Gene expression profiles in liver cancer  
Patent: WO 0229103-A 8 11-APR-2002;  
GENE LOGIC INC (US)

## FEATURES

Location/Qualifiers  
1..457  
/organism="Homo sapiens"

## source

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="EMBL/GenBank Accession No. AA004521"

## ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 457;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
DB 44 GCATGTTACAGGTAGAAAAGCC 65

## RESULT 3

AX880371/c AX880371 2093 bp DNA linear PAT 17-DEC-2003  
LOCUS AX880371 15276 from Patent EP1074617.  
DEFINITION Sequence 15276 from Patent EP1074617.  
ACCESSION AX880371  
VERSION AX880371.1 GI:40035107

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
Primers for synthesizing full-length cDNA and their use  
Patent: EP 1074617-A 15276 07-FEB-2001;  
Research Association for Biotechnology (JP)

## FEATURES

Location/Qualifiers  
1..2093  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
77..1534  
/note="unnamed protein product"

## CDS

/codon\_start=1  
/protein\_id="CAE90921.1"  
/db\_xref="GI:40035108"  
/translation="MNLGFTIGSFVLSATLPLGILMDRFGPRPRLVGSACFTASC  
TMALASRDVEALSPFLFSLNSLNGPGGICLTFSLTLPNMFNGLSTLMALMIGSYA  
KSATTFPGIKLIYDAGVAFVIMFTWSGLACLIPLNCTLNPIFAFPPEVNTYTKI  
KSLGLALDHKVTGDLFTYHTVMGRLSKAPSLSDGDFMSQDVHGTSNLPERS  
VPLRSLCSPTFLWSLTMGTQLRIIFYMAAVNMKLEYLVGTQDEHDTNQKQVAE  
TVGFVSSVEGAMQLCLLTCTPLIGVIMDWIKDCVADPTQCTGQVGHEDRGVATKIRP  
RYCKIKLTNAISAPTLTNLLVGGITCLINLHLQFTVFLVHTIVRGFFHSACGSL  
YAAVFPNSNHGTLTGLQSLISAVFALLOQPLPMAMVGPLKGEPPFWNLGLLLFSLLG  
LLPSYLFYTRARLQOEYAAANGWGPLKVLGSSEVTA"

## ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 2093;  
Best Local Similarity 100.0%; Pred. No. 2.8; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
DB 1625 GCATGTTACAGGTAGAAAAGCC 1604

## RESULT 4

BD158342/c

LOCUS BD158342

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD158342

VERSION BD158342.1 GI:27864100

KEYWORDS JP 2002191363-A/13185.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2093)

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004, Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 14:24:52 ; Search time 975.576 Seconds  
(without alignments)  
10128.688 Million cell updates/sec

Title: US-09-743-825-1  
Perfect score: 2326  
Sequence: 1 cgggctgaggggggcaaa.....agggagtgagaaaaaaa 2326

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3373863 seqs, 2124099041 residues  
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001s:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2326	100.0	2326	3	Aaz50442 Human PB3
2	2326	100.0	2326	9	Adb75506 Prostate
3	2214.6	95.2	2373	2	Aav49590 Human liv
4	2150	92.4	2284	7	Acc72139 Human NOV
5	2128.4	91.5	2233	7	Accd20413 DNA encod
6	2066.6	88.8	2093	4	Aah16350 Human CDN
7	1914	82.3	2105	7	Acc72138 Human NOV
8	1677	72.1	1677	2	Aav49591 Human liv
9	1613	69.3	3442	3	Aaz50443 Human PB3
10	684.2	29.4	1737	4	AAL05399 Human rep
11	684.2	29.4	1737	4	ABL98261 Human tes
12	684.2	29.4	3160	4	AAL05398 Human rep
13	684.2	29.4	3160	4	ABL98260 Human tes
14	684.2	29.4	7709	4	AAL05397 Human rep
15	684.2	29.4	7709	4	ABL98259 Human tes
16	684.2	29.4	143899	6	AAL38336 Genomic s
17	627.4	27.0	632	4	AAH07095 Human CDN
18	616.2	26.5	2037	7	Aad54747 Human CGD
19	612.8	26.3	2934	4	AAD10330 Human hae
20	609.6	26.2	2934	4	AAD10329 Human hae
21	540	23.2	2643	4	AAK94522 Human ful
22	503	21.6	2589	4	AAK94430 Human ful
23	492.2	21.2	579	4	AAH12098 Human CDN

C	24	464.2	20.0	2490	4	AAD05103	Aad05103 Human sec
C	25	426.2	18.3	2078	6	ABX34031	Abx34031 Human can
C	26	423.6	18.2	464	6	ABL64481	AbL64481 Stomach c
C	27	423.6	18.2	464	6	ABN94041	Abn94041 Gene #539
	28	403.4	17.3	410	8	ACH21519	Ach21519 Human adu
	29	376.8	16.2	2349	4	AAK94539	AAK94539 Human ful
	30	359.6	15.5	416	8	ACH20959	Ach20959 Human adu
	31	306	13.2	859	4	AAD10328	Aad10328 Human hae
	32	289	12.4	630	4	AAK93483	AAK93483 Human CDN
	33	289	12.4	630	4	AAK92221	AAK92221 Human CDN
C	34	281.6	12.1	856	4	AAL01556	Aal01556 Human rep
C	35	281.6	12.1	856	4	ABL96993	AbL96993 Human tes
	36	277.4	11.9	1359	3	AAZ56705	Aaz56705 Human tra
	37	276.6	11.9	3771	5	AAZ74359	Aaz74359 DNA encod
C	38	271.6	11.7	457	5	ABN93510	Abn93510 Gene #8 u
	39	265.4	11.4	1479	5	AAZ79359	Aaz79359 DNA encod
	40	252.6	10.9	613	4	AAK92120	AAK92120 Human CDN
	41	252.6	10.9	613	4	AAK93384	AAK93384 Human CDN
	42	245.8	10.6	1518	6	ABL90731	AbL90731 Human pol
	43	230.4	9.9	4121	7	ACC46528	Acc46528 Human dit
	44	194.2	8.3	800	4	AAK93502	AAK93502 Human CDN
	45	194.2	8.3	800	4	AAK92242	AAK92242 Human CDN

## ALIGNMENTS

RESULT 1  
AAZ50442  
ID AAZ50442 standard; cDNA; 2326 BP.

XX AAZ50442;

DT 18-MAY-2000 (first entry)

XX Human PB39 cDNA, a gene dysregulated in prostate cancer.

XX PB39; human; prostate cancer; PC; chromosome 11p11.1-11.2; cancer;  
prostate epithelium; splicing mechanism; early diagnosis; progression;  
precancerous cell; aggressive prostate carcinoma; metastatic potential;  
non-neoplastic prostate disease; expressed sequence tag; EST; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT 5'UTR 1...76

FT misc\_signal /tag= a

FT /tag= b

FT label= Kozak\_consensus\_sequence

FT CDS 77..1756

FT /tag= c

FT /product= "Human PB39 protein"

FT 3'UTR /note= "Over-expressed in prostate cancer"

FT 1757..2326

FT /tag= d

XX WO200005376-A1.

XX 03-FEB-2000.

XX 23-JUL-1999; 99WO-US016831.

XX 24-JUL-1998; 98US-0094137P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Chuquai RF, Cole KA, Liotta LA;

XX WPI; 2000-182700/16.

XX P-PSDB; AAY44897.

XX Novel gene which is dysregulated in prostate cancer useful for diagnosing





QY 781 GCTGAGTGGGCTGGCCCTGGACCAAGAGTGACAGGTGACCTCTCTACACCCATGTGAC 840  
 Db 781 GCTGAGTGGGCTGGCCCTGGACCAAGAGTGACAGGTGACCTCTCTACACCCATGTGAC 840  
 QY 841 CACCATGGCCAGAGGCTCAGCCAGAGGCCCCCAGCCTGGAGGACGGTTGGGATGCGCTT 900  
 Db 841 CACCATGGCCAGAGGCTCAGCCAGAGGCCCCCAGCCTGGAGGACGGTTGGGATGCGCTT 900  
 QY 901 CATGTACCCGAGGATGTTGGGGGACCTCAGAAAACTCTCTGAGAGGTCTGTCCCTT 960  
 Db 901 CATGTACCCGAGGATGTTGGGGGACCTCAGAAAACTCTCTGAGAGGTCTGTCCCTT 960  
 QY 961 AGCGAAGAGCCTCTGCTCCCCACCTTCTGTGGAGCCTCTCACCATGGGATGACCCCA 1020  
 Db 961 AGCGAAGAGCCTCTGCTCCCCACCTTCTGTGGAGCCTCTCACCATGGGATGACCCCA 1020  
 QY 1021 GCTGCGGATCATCTTCTACATGGCTGTGTGAACAAGATGTGGAGTACCTTGTGACTGG 1080  
 Db 1021 GCTGCGGATCATCTTCTACATGGCTGTGTGAACAAGATGTGGAGTACCTTGTGACTGG 1080  
 QY 1081 TGGCCAGAGCATGAGCAAAATGAACAGCAACAAAGGTGGCAGAGACAGTTGGGTTCTA 1140  
 Db 1081 TGGCCAGAGCATGAGCAAAATGAACAGCAACAAAGGTGGCAGAGACAGTTGGGTTCTA 1140  
 QY 1141 CTCTCTCGTCTTGGGGGCGATGACGCTGTGTGGCTTCTCACCTGCCCTCATTTGGCTA 1200  
 Db 1141 CTCTCTCGTCTTGGGGGCGATGACGCTGTGTGGCTTCTCACCTGCCCTCATTTGGCTA 1200  
 QY 1201 CATCATGAGCTGGCGGATCAAGGACTGTGGTGACGCCCCCACTCAGGGCAGCTGTCTCGG 1260  
 Db 1201 CATCATGAGCTGGCGGATCAAGGACTGTGGTGACGCCCCCACTCAGGGCAGCTGTCTCGG 1260  
 QY 1261 AGATGCCAGGAGCGGGTTGCTACCAATCCATCAGACACACGCTACTGCAAGATCCAAAA 1320  
 Db 1261 AGATGCCAGGAGCGGGTTGCTACCAATCCATCAGACACACGCTACTGCAAGATCCAAAA 1320  
 QY 1321 GCTCACCAGTGCATCAGTGCCTTCCCTGACCAACCTGCTGTGGGTTTGGCAT 1380  
 Db 1321 GCTCACCAGTGCATCAGTGCCTTCCCTGACCAACCTGCTGTGGGTTTGGCAT 1380  
 QY 1381 CACCTGTCTCATCAACAATTACACTCCAGTTTGTGACCTTTGTCTGACACCATTTGT 1440  
 Db 1381 CACCTGTCTCATCAACAATTACACTCCAGTTTGTGACCTTTGTCTGACACCATTTGT 1440  
 QY 1441 TCGAGGTTTCTCAGCTGAGCTGTGGAGTCTATGCTGACAGTGTTCCTGCAACCA 1500  
 Db 1441 TCGAGGTTTCTCAGCTGAGCTGTGGAGTCTATGCTGACAGTGTTCCTGCAACCA 1500  
 QY 1501 CTTTGGGAGCTGACAGGCTGCACTCCCTCATCAGTGTGTGCTGCTTGTTCAGCA 1560  
 Db 1501 CTTTGGGAGCTGACAGGCTGCACTCCCTCATCAGTGTGTGCTGCTTGTTCAGCA 1560  
 QY 1561 GCCACTTTTTCATGGGATGGTGGGACCCCTGAAAGAGAGGCCCTTCTGGGTGAATCTGGG 1620  
 Db 1561 GCCACTTTTTCATGGGATGGTGGGACCCCTGAAAGAGAGGCCCTTCTGGGTGAATCTGGG 1620  
 QY 1621 CCTCTGTCTATCTCAGCTGCGGATTTCTGTGCTTCTTACCTCTCTATACCGTGC 1680  
 Db 1621 CCTCTGTCTATCTCAGCTGCGGATTTCTGTGCTTCTTACCTCTCTATACCGTGC 1680  
 QY 1681 CCGGCTCCAGCAGGAGTACGCGCCCAATGGGATGGGCGCCACTGAAAGGTGCTTAGCGGCTC 1740  
 Db 1681 CCGGCTCCAGCAGGAGTACGCGCCCAATGGGATGGGCGCCACTGAAAGGTGCTTAGCGGCTC 1740  
 QY 1741 TGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTTGATGACAGGCAATCAAGGCTG 1800  
 Db 1741 TGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTTGATGACAGGCAATCAAGGCTG 1800  
 QY 1801 AGCAACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTGTAACTGACATAGAGCCATG 1860  
 Db 1801 AGCAACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTGTAACTGACATAGAGCCATG 1860

QY 1861 GCCGTAGATTATTAATAACCAAGAGAGTCTTATTTTTGTAAAGACTGCAAAAAGGAGGA 1920  
 Db 1861 GCCGTAGATTATTAATAACCAAGAGAGTCTTATTTTTGTAAAGACTGCAAAAAGGAGGA 1920  
 QY 1921 AAAAAAACCTTTCAAAAAACGCCCCCTAAAGTCAACGCTCCATTGACTGAAGACAGTCCCTAT 1980  
 Db 1921 AAAAAAACCTTTCAAAAAACGCCCCCTAAAGTCAACGCTCCATTGACTGAAGACAGTCCCTAT 1980  
 QY 1981 CCTAGAGGGTTGAGCTTCTTCTCTCTGGTTGGAGGAGACCAAGGTCCTCTTATCT 2040  
 Db 1981 CCTAGAGGGTTGAGCTTCTTCTCTCTGGTTGGAGGAGACCAAGGTCCTCTTATCT 2040  
 QY 2041 CTTTCTAGCGTCTGCGCTCCCTCTGCTACCTCTTGGGGGATCGGCAACAGGCTACCCCTGA 2100  
 Db 2041 CTTTCTAGCGTCTGCGCTCCCTCTGCTACCTCTTGGGGGATCGGCAACAGGCTACCCCTGA 2100  
 QY 2101 GGTCCCATGTGCCATGAGTGTGCAACAATGATGTCTGTGTATGTGTGAATGTGAG 2160  
 Db 2101 GGTCCCATGTGCCATGAGTGTGCAACAATGATGTCTGTGTATGTGTGAATGTGAG 2160  
 QY 2161 AAAAAACAGCCCTCTTTCAGAAAGGAAAGGGGCTGAGGTGCCAGCTGTCTCTGGGTT 2220  
 Db 2161 AAAAAACAGCCCTCTTTCAGAAAGGAAAGGGGCTGAGGTGCCAGCTGTCTCTGGGTT 2220  
 QY 2221 AGGGGTTGGGGTCTGGGCCCTTCCAGGCGCCAGGAGGAGGCTTCCCTCTCTGGTGTCT 2280  
 Db 2221 AGGGGTTGGGGTCTGGGCCCTTCCAGGCGCCAGGAGGAGGCTTCCCTCTCTGGTGTCT 2280  
 QY 2281 GCTTGCAGTCTTAGAGGAAATAAAAAAGGGAAGTGAAGAAAAA 2326  
 Db 2281 GCTTGCAGTCTTAGAGGAAATAAAAAAGGGAAGTGAAGAAAAA 2326

RESULT 3  
 AAV49590  
 ID AAV49590 standard; cDNA to mRNA; 2373 BP.  
 XX AC AAV49590;  
 XX DT 21-OCT-1998 (first entry)  
 XX Human liver cell clone HP10302 cDNA #1.  
 DE Transmembrane domain; human; nutrition; cytokine; cell proliferation;  
 KW differentiation; immune system; stimulator; suppressor; regulator;  
 KW hematopoiesis; activin; inhibitor; chemotactic; chemokinetic; receptor;  
 KW haemostatic; thrombolytic; ligand; anti-inflammatory; tumour; ds.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT CDS 134..1813  
 FT /\*tag= a  
 FT /product= "transmembrane domain containing protein"  
 XX WO9821328-A2.  
 XX PD 22-MAY-1998.  
 XX PF 07-NOV-1997; 97WO-JP004056.  
 XX PR 13-NOV-1996; 96JP-00301429.  
 XX PA (SAGA ) SAGAMI CHEM RES CENTRE.  
 XX PA (PROT-) PROTEGENE INC.  
 XX PI Kato S, Sekine S, Yamaguchi T, Kobayashi M;  
 DR WPI; 1998-297932/26.  
 DR P-FSDB; AAW64554.  
 XX Human protein having transmembrane domain - useful for, e.g. research and  
 PT nutrition.

XX Claim 4; Page 166-169; 205pp; English.  
 PS AAV49550-V49599 are cDNA sequences which encode human proteins containing  
 CC a transmembrane domain. These proteins can be used for, e.g. research and  
 CC nutrition, and may have cytokine and cell proliferation/differentiation,  
 CC immune stimulating/suppressing, haematopoiesis regulating, tissue growth,  
 CC activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic,  
 CC receptor/ligand, anti-inflammatory or tumour inhibition activity  
 XX  
 SQ Sequence 2373 BP; 469 A; 716 C; 645 G; 543 T; 0 U; 0 Other;  
  
 Query Match 95.2%; Score 2214.6; DB 2; Length 2373;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 2243; Conservative 0; Mismatches 9; Indels 3; Gaps 2;  
  
 QY 64 GCTCGGGGCGAGATGGCCCGCCCGCTGCAACAGGCGGTACCGAGGCGGTGGTGGATGGC 123  
 DB 121 GCTCGGGGCGAGATGGCCCGCCCGCTGCAACAGGCGGTACCGAGGCGGTGGTGGATGGC 180  
  
 QY 124 CTGACGGCTGTGCTGGAGAACCTCTTCTCTGTCTGTACTCTCTGGGCTGGGGCTCCCT 183  
 DB 181 CTGACGGCTGTGCTGGAGAACCTCTTCTCTGTCTGTACTCTCTGGGCTGGGGCTCCCT 240  
  
 QY 184 GTTGATCATTTCTGAAGAACGAGGGCTTTCTATTCCAGCAGCTGCCAGCTGAGAGCAGCAC 243  
 DB 241 GTTGATCATTTCTGAAGAACGAGGGCTTTCTATTCCAGCAGCTGCCAGCTGAGAGCAGCAC 300  
  
 QY 244 CAACACACCGAGGATGAGAGGCGAGGTGGCGAGCTGTGACAGCAGGACGAGATGCT 303  
 DB 301 CAACACACCGAGGATGAGAGGCGAGGTGGCGAGCTGTGACAGCAGGACGAGATGCT 360  
  
 QY 304 CAACCTGGGCTTCAACATTTGGTTCTTCTGTCTGCTCAGCGCCACCCCTCCCTGGGGAT 363  
 DB 361 CAACCTGGGCTTCAACATTTGGTTCTTCTGTCTGCTCAGCGCCACCCCTCCCTGGGGAT 420  
  
 QY 364 CTCTATGAGACCGCTTTGGCCCGGACCCCGTGGGCTGGTTGGAGCTGCTGTCTTCACTGC 423  
 DB 421 CTCTATGAGACCGCTTTGGCCCGGACCCCGTGGGCTGGTTGGAGCTGCTGTCTTCACTGC 480  
  
 QY 424 GTCTGACACCTCATGGCCCTGGCCCTCCGGAGCTGGAAGCTCTGTCTCGTTGATATT 483  
 DB 481 GTCTGACACCTCATGGCCCTGGCCCTCCGGAGCTGGAAGCTCTGTCTCGTTGATATT 540  
  
 QY 484 CTGGGCGCTGCTCCGTAATGGCTTTGTGTCATCTGCTTAACTGCTTCACTACGCT 543  
 DB 541 CTGGGCGCTGCTCCGTAATGGCTTTGTGTCATCTGCTTAACTGCTTCACTACGCT 600  
  
 QY 544 GCCCAACATGTTGGGAACCTGGCTCCAGCTTAATGGCCCTCATGATTTGGCTTTTACGC 603  
 DB 601 GCCCAACATGTTGGGAACCTGGCTCCAGCTTAATGGCCCTCATGATTTGGCTTTTACGC 660  
  
 QY 604 CTCTTCTGCATTTACGTTCCAGGAATCAAGCTGATCTACGATGCGGCTGGCTTCGT 663  
 DB 661 CTCTTCTGCATTTACGTTCCAGGAATCAAGCTGATCTACGATGCGGCTGGCTTCGT 720  
  
 QY 664 GGTTCATCATGTTCACTGTGCTGGCTGGCTGGCTTTATCTTCTGAACTGACCCCTCAA 723  
 DB 721 GGTTCATCATGTTCACTGTGCTGGCTGGCTGGCTTTATCTTCTGAACTGACCCCTCAA 780  
  
 QY 724 CTGGCCCATCGAAGCCTTTCTGCCCCCTGAGGAAGTCAATTACAGGAAGATCAAGCT 783  
 DB 781 CTGGCCCATCGAAGCCTTTCTGCCCCCTGAGGAAGTCAATTACAGGAAGATCAAGCT 840  
  
 QY 784 GAGTGGGCTGGCCCTGGACCAACAGGTGACGTGACCTTCTTACACCCATGTGACCCAC 843  
 DB 841 GAGTGGGCTGGCCCTGGACCAACAGGTGACGTGACCTTCTTACACCCATGTGACCCAC 900  
  
 QY 844 CATGGCCACAGAGGCTCAGCCAGAGGCCCCCGAGCTGGAGGACGGTTCGGATGCCCTTCA 903  
 DB 901 CATGGCCACAGAGGCTCAGCCAGAGGCCCCCGAGCTGGAGGACGGTTCGGATGCCCTTCA 960  
  
 QY 904 GTACCCCGAGGATGTTTCGGGGCACCTTCAGAAAACCTTCTCTGAGAGGCTGTGCCCTTACG 963

DB 961 GTACCCCGAGGATGTTGGGGCACCTCAGAAAACCTTCTCTGAGAGGCTGTGCCCTTACG 1020  
 QY 964 CAAAGAGCTTCTCTCTCCCCCACCCTTCTCTGAGAGGCTTCTCACCATGGGATGACCCAGCT 1023  
 DB 1021 CAAAGAGCTTCTCTCTCCCCCACCCTTCTCTGAGAGGCTTCTCACCATGGGATGACCCAGCT 1080  
 QY 1024 GGGGATCATTTCTACATGGCTGTGTGAAACAAGATGCTGGAGTACCTTGTGACTGGTGG 1083  
 DB 1081 GGGGATCATTTCTACATGGCTGTGTGAAACAAGATGCTGGAGTACCTTGTGACTGGTGG 1140  
 QY 1084 CAGAGCATGAGACAAATGAAACAGCAACAAAGGTGGCAGAGACAGTGGGTTCATCTC 1143  
 DB 1141 CAGAGCATGAGACAAATGAAACAGCAACAAAGGTGGCAGAGACAGTGGGTTCATCTC 1200  
 QY 1144 CTCCGCTCTTCTGGGGCATGAGCTGTGTGCTTCTCAGCTGCCCTCATTTGGCTACAT 1203  
 DB 1201 CTCCGCTCTTCTGGGGCATGAGCTGTGTGCTTCTCAGCTGCCCTCATTTGGCTACAT 1260  
 QY 1204 CATGAGCTGGCGGATCAAGGACTGGGTGAGCGCCCCCAACTCAGGGGCACTGTCTCGGAGA 1263  
 DB 1261 CATGAGCTGGCGGATCAAGGACTGGGTGAGCGCCCCCAACTCAGGGGCACTGTCTCGGAGA 1320  
 QY 1264 TGCCAGGAGCGGGTGTCTACAAATCCATCAGACACACGCTACTGCAAGATCCAAAGCT 1323  
 DB 1321 TGCCAGGAGCGGGTGTCTACAAATCCATCAGACACACGCTACTGCAAGATCCAAAGCT 1380  
 QY 1324 CACCAATGCCATCAGTGCCTTCACTCAGCTCCAGTTCGTGACCTTGTCTGACACACCATTTCTG 1383  
 DB 1381 CACCAATGCCATCAGTGCCTTCACTCAGCTCCAGTTCGTGACCTTGTCTGACACACCATTTCTG 1440  
 QY 1384 CTGTCTCATCAACAACTTACACCTCCAGTTCGTGACCTTGTCTGACACACCATTTCTG 1443  
 DB 1441 CTGTCTCATCAACAACTTACACCTCCAGTTCGTGACCTTGTCTGACACACCATTTCTG 1500  
 QY 1444 AGGTTTCTTCCATCTCAGCTGTGGAGTCTCTATGCTGAGTGTTCCTTCCATCCAAACCTT 1503  
 DB 1501 AGGTTTCTTCCATCTCAGCTGTGGAGTCTCTATGCTGAGTGTTCCTTCCATCCAAACCTT 1560  
 QY 1504 TGGGAGCTGACAGGCGCTGAGTCCCTCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1563  
 DB 1561 TGGGAGCTGACAGGCGCTGAGTCCCTCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 QY 1564 ACTTTTCTATGGGATGGTGGGACCCCTGAAAGGAGAGCCCTTCTGGGTGAACTGCGGCT 1623  
 DB 1621 ACTTTTCTATGGGATGGTGGGACCCCTGAAAGGAGAGCCCTTCTGGGTGAACTGCGGCT 1680  
 QY 1624 CTTGCTATTTCTCATCTCTGGGATCTCTGTGCTTCTTACCTTCTTCTATTACCGTCCCG 1683  
 DB 1681 CTTGCTATTTCTCATCTCTGGGATCTCTGTGCTTCTTACCTTCTTCTATTACCGTCCCG 1740  
 QY 1684 GCTCAGCAGGATGAGCGCGCAATGGGATGGGCGCCCATGAGGTGCTTTAGGGCTCTGA 1743  
 DB 1741 GCTCAGCAGGATGAGCGCGCAATGGGATGGGCGCCCATGAGGTGCTTTAGGGCTCTGA 1800  
 QY 1744 GGTGACCGCATAGACTTCTCAGACCAAGGAGCTGGATGACAGGCAATCAAGGCTGAGC 1803  
 DB 1801 GGTGACCGCATAGACTTCTCAGACCAAGGAGCTGGATGACAGGCAATCAAGGCTGAGC 1860  
 QY 1804 AACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTTGTAACTGTAACATGACATAGGCCAT 1863  
 DB 1861 AACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTTGTAACTGTAACATGACATAGGCCAT 1920  
 QY 1864 GTAGATTTTAAATACCAAGAGAGTCTTATTTTGTAAAGTCTCCAAAGAGGAGG - AAA 1922  
 DB 1921 GTAGATTTTAAATACCAAGAGAGTCTTATTTTGTAAAGTCTCCAAAGAGGAGGAGAAA 1980  
 QY 1923 AAAAAACCTTCAAAAAAGCGCCCTAAGTCAACGCTCCATTTGACTGAAGACAGTCCCTATCC 1982  
 DB 1981 AAAAAACCTTCAAAAAAGCGCCCTAAGTCAACGCTCCATTTGACTGAAGACAGTCCCTATCC 2040  
 QY 1983 TAGAGGGTGTAGCTTTTCTTCTCTGGTGGAGGAGACAGGGTGTCTTATCTTCTTC 2042









1201 CATCATGGAGTGGCGGATCAAGGACTGCGTGGAGCGCCCAACTCAGGCGACTGCTCTCG 1260  
 1108 CATCATGGAGTGGCGGATCAAGGACTGCGTGGAGCGCCCAACTCAGGCGACTGCTCTCG 1167  
 1261 AGATGCGAGGAGCGGGTGTCTACCAAAATCCATCAGACGACGCTACTGCAAGATCAAAA 1320  
 1168 AGATGCGAGGAGCGGGTGTCTACCAAAATCCATCAGACGACGCTACTGCAAGATCAAAA 1227  
 1321 GCTCACAATGCGATCAGTGCCTTCCACCTGACCAACCTGCTGCTGCTGCTGCTGCTG 1380  
 1228 GCTCACAATGCGATCAGTGCCTTCCACCTGACCAACCTGCTGCTGCTGCTGCTGCTG 1287  
 1381 CACCTGCTCATCAACAACCTTACACCTCCAGTTTGTGACCTTGTGCTGCTGCTGCTGCTG 1440  
 1288 CACCTGCTCATCAACAACCTTACACCTCCAGTTTGTGACCTTGTGCTGCTGCTGCTGCTG 1347  
 1441 TCGAGTTTCTTCCACTCAGCTGCTGCGAGTCTCTATGCTGCGAGTGTCCATCCAAACA 1500  
 1348 TCGAGTTTCTTCCACTCAGCTGCTGCGAGTCTCTATGCTGCGAGTGTCCATCCAAACA 1407  
 1501 CTTTGGAGCGCTGACAGGCTGCGAGTCCCTCATCAGTGTGCTGCTGCTGCTGCTGCTG 1560  
 1408 CTTTGGAGCGCTGACAGGCTGCGAGTCCCTCATCAGTGTGCTGCTGCTGCTGCTGCTG 1467  
 1561 GCCACTTTTTCATGGCGATGGTGGACCCCTGAAAGGAGAGCCCTTCTGGGTGCAATCTGGG 1620  
 1468 GCCACTTTTTCATGGCGATGGTGGACCCCTGAAAGGAGAGCCCTTCTGGGTGCAATCTGGG 1527  
 1621 CTTCTGCTATTTCTCCTCTGCGGATCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
 1528 CTTCTGCTATTTCTCCTCTGCGGATCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1587  
 1681 CCGGCTCCAGCAGGAGTACGCGCCCAATGGGATGGGCGCCACTGAGGCTGCTGAGCGCTC 1740  
 1588 CCGGCTCCAGCAGGAGTACGCGCCCAATGGGATGGGCGCCACTGAGGCTGCTGAGCGCTC 1647  
 1741 TGAGGTGACCGCATAGACTTCTCAGACCAAGGAGCCTGGATGACAGGCAATCAAGGCTG 1800  
 1648 TGAGGTGACCGCATAGACTTCTCAGACCAAGGAGCCTGGATGACAGGCAATCAAGGCTG 1707  
 1801 AGCAACCAAAAGAGTGGCGCCCATATGCTTTTCTACCTGTATACATGCAATGAGGCGATG 1860  
 1708 AGCAACCAAAAGAGTGGCGCCCATATGCTTTTCTACCTGTATACATGCAATGAGGCGATG 1767  
 1861 GCGGTAGATTATAATACCAAGAGAGTCTATTTTCTAAGACTGCAAAAGGAGCA 1920  
 1768 GCGGTAGATTATAATACCAAGAGAGTCTATTTTCTAAGACTGCAAAAGGAGCA 1827  
 1921 AAAAAAACCTTCAAAAACGCCCCCTAAGTCAACGCTCCATTGACTGAAGACAGTCCCTAT 1980  
 1828 AAAAAAACCTTCAAAAACGCCCCCTAAGTCAACGCTCCATTGACTGAAGACAGTCCCTAT 1887  
 1981 CTTAGAGGGGTGAGCTTTCTCTCTGCTGGTGGAGAGACAGGGGTGCTCTTATCT 2040  
 1888 CTTAGAGGGGTGAGCTTTCTCTCTGCTGGTGGAGAGACAGGGGTGCTCTTATCT 1947  
 2041 CTTTCTAGCGGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
 1948 CTTTCTAGCGGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2007  
 2101 GGTGCCATGTGCGCATGAGTGTGCAACAATGCAATGCTGTGTATGTGTGAATGTGAG 2160  
 2008 GGTGCCATGTGCGCATGAGTGTGCAACAATGCAATGCTGTGTATGTGTGAATGTGAG 2067  
 2161 AAAAAACAGCCCTCTTTCAGAGGAAAGGGGCTGAGGTGCGAGCTGTGCTGCTGCTGCTG 2220  
 2068 AAAAAACAGCCCTCTTTCAGAGGAAAGGGGCTGAGGTGCGAGCTGTGCTGCTGCTGCTG 2127  
 2221 AGGGGTGGGGTGGCGCCCTTCCAGGCGCAGGAAGGAGGTTCCTCTCTGCTGCTGCTGCT 2280  
 2128 AGGGGTGGGGTGGCGCCCTTCCAGGCGCAGGAAGGAGGTTCCTCTCTGCTGCTGCTGCT 2187

QY 2281 GCTTGCAGCTCTTAGAGGAAATAAAAAGGAAAGTGAAGAAAAA 2326  
 DB 2188 GCTTGCAGCTCTTAGAGGAAATAAAAAGGAAAGTGAAGAAAAA 2233

RESULT 6  
 AAH16350  
 ID AAH16350 standard; cDNA; 2093 BP.  
 AC AAH16350;  
 XX 26-JUN-2001 (first entry)  
 DT Human cDNA sequence SEQ ID NO:15276.  
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 KW Homo sapiens.  
 OS EP1074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-00116126.  
 XX 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX (HELI-) HELIX RES INST.  
 XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 DR WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 15276; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any special methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

Sequence 2093 BP; 429 A; 623 C; 548 G; 493 T; 0 U; 0 Other;

Query Match 88.8%; Score 2066.6; DB 4; Length 2093;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 2094; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 223 GTGCCAGCTGAGACGACCAACACACCCAGAGATGAGCAGCGCAGGTGGCCAGGCTG 282  
DB 1 GTGCCAGCTGAGACGACCAACACACCCAGAGATGAGCAGCGCAGGTGGCCAGGCTG 60  
QY 283 TGACCAGAGCAGAGATGCTCAACCTGGGCTTACCAATGGTTGCTTCTCGTCTCAGCGC 342  
DB 61 TGACCAGAGCAGAGATGCTCAACCTGGGCTTACCAATGGTTGCTTCTCGTCTCAGCGC 120  
QY 343 CACCACTGCGACCTGAGGATCCTCATGACCGCTTTTGGCCCCGACCCGCTGCGGCTGGT 402  
DB 121 CACCACTGCGACCTGAGGATCCTCATGACCGCTTTTGGCCCCGACCCGCTGCGGCTGGT 180  
QY 403 TGGCAGTGCCTGCTTCACTGCGCTGCTGACCCCTCATGCGCTGCGGCTCCCGGAGCTGGA 462  
DB 181 TGGCAGTGCCTGCTTCACTGCGCTGCTGACCCCTCATGCGCTGCGGCTCCCGGAGCTGGA 240  
QY 463 AGCTCTGCTCCGTTGATATTCTTGGCGCTGTCCCTGAATGGCTTTGGTGGCATCTGCCT 522  
DB 241 AGCTCTGCTCCGTTGATATTCTTGGCGCTGTCCCTGAATGGCTTTGGTGGCATCTGCCT 300  
QY 523 AACGTTCACTTCACTCAGCTGCCCAACATGTTTGGGAACCTGCGCTCCAGTTAATGGC 582  
DB 301 AACGTTCACTTCACTCAGCTGCCCAACATGTTTGGGAACCTGCGCTCCAGTTAATGGC 360  
QY 583 CCTCATGATTGGCTCTTACGGCTCTTCTGCCATTACGTTCCAGGAATCAAGCTGATCTA 642  
DB 361 CCTCATGATTGGCTCTTACGGCTCTTCTGCCATTACGTTCCAGGAATCAAGCTGATCTA 420  
QY 643 CGATGCCGCTGTGGCTTCTGCTGGTCACTCATGTTTACCTGGCTGGCTGCTGCTTAT 702  
DB 421 CGATGCCGCTGTGGCTTCTGCTGGTCACTCATGTTTACCTGGCTGGCTGCTGCTTAT 480  
QY 703 CTTTCTGAATCTGACCTCACTGCGCCATCGAAGCCTTCTGCGCCCTGAGGAAGTCAA 762  
DB 481 CTTTCTGAATCTGACCTCACTGCGCCATCGAAGCCTTCTGCGCCCTGAGGAAGTCAA 540  
QY 763 TTACACGAGAGATCAAGCTGAGTGGCTGGCTGGACCAACAGGTGACAGTGCCT 822  
DB 541 TTACACGAGAGATCAAGCTGAGTGGCTGGCTGGACCAACAGGTGACAGTGCCT 600  
QY 823 CTTTCTACCCATGTGACCCACCATGGGCGCAGAGGCTCAGCCAGAGGGCCCCCAGCCTGGA 882  
DB 601 CTTTCTACCCATGTGACCCACCATGGGCGCAGAGGCTCAGCCAGAGGGCCCCCAGCCTGGA 660  
QY 883 GGAAGTTCGGATGCTTCACTGTCACCCAGAGTGTTCGGGCACTCAGAAAACTTCC 942  
DB 661 GGAAGTTCGGATGCTTCACTGTCACCCAGAGTGTTCGGGCACTCAGAAAACTTCC 720  
QY 943 TGAGAGGTCTGCTCCCTTACGCAAGAGCTGTGCTCCCGCACTTTCTGTTGGAGCTCCT 1002  
DB 721 TGAGAGGTCTGCTCCCTTACGCAAGAGCTGTGCTCCCGCACTTTCTGTTGGAGCTCCT 780  
QY 1003 CACCATGGGCATGACCCAGCTCGGATCATTTCTACATGGCTGTGTGAACAAGATGCT 1062  
DB 781 CACCATGGGCATGACCCAGCTCGGATCATTTCTACATGGCTGTGTGAACAAGATGCT 840  
QY 1063 GGAATACCTTGTACTGTGGTGGCCAGGATGAGACAAATGAACAGCAACAAAGGTGGC 1122  
DB 841 GGAATACCTTGTACTGTGGTGGCCAGGATGAGACAAATGAACAGCAACAAAGGTGGC 900  
QY 1123 AGAGACAGTTGGGTTCTACTCTCGTCTTTCGGGGCCATGACAGTGTGTGCTTCTCAC 1182  
DB 901 AGAGACAGTTGGGTTCTACTCTCGTCTTTCGGGGCCATGACAGTGTGTGCTTCTCAC 960  
QY 1183 CTTGCCCCCTCATTTGGCTACATCATGACTGGCGGATCAAGACTCGCTGGAGCCCCAAC 1242  
DB 961 CTTGCCCCCTCATTTGGCTACATCATGACTGGCGGATCAAGACTCGCTGGAGCCCCAAC 1020  
QY 1243 TCAGGGCACTGTCTCGGAGTGCACGGACGGGTTGCTACCAATCCATCAGACCAG 1302  
DB 1021 TCAGGGCACTGTCTCGGAGTGCACGGACGGGTTGCTACCAATCCATCAGACCAG 1080

QY 1303 CTACTGCAAGATCAAAGGCTCACAAATGCCATGAGTGCCTTCAACCTGCACCAACCTGCT 1362  
DB 1081 CTACTGCAAGATCAAAGGCTCACAAATGCCATGAGTGCCTTCAACCTGCACCAACCTGCT 1140  
QY 1363 GCTTGTGGGTTTGGCATCACCTGTCTCATCAACAACCTTACACCTCCAGTTTGTGACCTT 1422  
DB 1141 GCTTGTGGGTTTGGCATCACCTGTCTCATCAACAACCTTACACCTCCAGTTTGTGACCTT 1200  
QY 1423 TGTCTGCACACCAATTTGTGAGGTTTCTTCCACTGAGCTGTGGAGTCTCTATGCTGC 1482  
DB 1201 TGTCTGCACACCAATTTGTGAGGTTTCTTCCACTGAGCTGTGGAGTCTCTATGCTGC 1260  
QY 1483 AGTGTCTCCCATCCAAACACTTTTGGAGCGTGCAGAGGCTGCAGTCCCTCATCAGTGTGT 1542  
DB 1261 AGTGTCTCCCATCCAAACACTTTTGGAGCGTGCAGAGGCTGCAGTCCCTCATCAGTGTGT 1320  
QY 1543 GTTCGCTTGTCTCAGCAGCCACTTTTTCATGGCGATGGTGGGACCCCTGAAAGGAGAGCC 1602  
DB 1321 GTTCGCTTGTCTCAGCAGCCACTTTTTCATGGCGATGGTGGGACCCCTGAAAGGAGAGCC 1380  
QY 1603 CTTCTGGGTGATCTGGGCTCCTGCTATTCTCACTCTGGGATTCCTGTTGCCCTTCCTA 1662  
DB 1381 CTTCTGGGTGATCTGGGCTCCTGCTATTCTCACTCTGGGATTCCTGTTGCCCTTCCTA 1440  
QY 1663 CTTCTTATTATACCGTGCCTGGCTCCAGCAGAGGTAAGCCGCCAATGGGATGGGCCCCACT 1722  
DB 1441 CTTCTTATTATACCGTGCCTGGCTCCAGCAGAGGTAAGCCGCCAATGGGATGGGCCCCACT 1500  
QY 1723 GAAGGTCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGATG 1782  
DB 1501 GAAGGTCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGATG 1560  
QY 1783 ACAGGCAATCAAGGCTTGAGCAACAAAGAGGTGCCCATATATGGCTTTCTACTGTGAA 1842  
DB 1561 ACAGGCAATCAAGGCTTGAGCAACAAAGAGGTGCCCATATATGGCTTTCTACTGTGAA 1620  
QY 1843 CATGCACATAGAGCATGGCGTAGATTATAAATACCAAGAGAAAGTCTATTTTGTAA 1902  
DB 1621 CATGCACATAGAGCATGGCGTAGATTATAAATACCAAGAGAAAGTCTATTTTGTAA 1680  
QY 1903 AGACTGCAAAAGAGGAAACCAACCTTCAAAAACGCCCCCTAAGTCAACGCTCCATG 1962  
DB 1681 AGACTGCAAAAGAGGAAACCAACCTTCAAAAACGCCCCCTAAGTCAACGCTCCATG 1740  
QY 1963 ACTCAAGACAGTCCCTATCTTAGAGGGGTGAGCTTTCTCTCTCTGGGTTGGAGGAGA 2022  
DB 1741 ACTCAAGACAGTCCCTATCTTAGAGGGGTGAGCTTTCTCTCTCTGGGTTGGAGGAGA 1800  
QY 2023 CCAGGGTGCCTTCTTATCTCTTAGCGGTCTGCTCTCTGTTACCTCTTTGGGGGATCGG 2082  
DB 1801 CCAGGGTGCCTTCTTATCTCTTCTTAGCGGTCTGCTCTCTGTTACCTCTTTGGGGGATCGG 1860  
QY 2083 CAAAACAGGCTACCCCTGAGGTCCCATGTGCATGAGTGTGCACAACTGCAATGTGCTG 2142  
DB 1861 CAAAACAGGCTACCCCTGAGGTCCCATGTGCATGAGTGTGCACA - CATGCAATGTGCTG 1918  
QY 2143 TGTATGTGTAATGTGAGAAACACAGCCCTCTTTCAGAGGAAAGGGGCTTGAGGTG 2202  
DB 1919 TGTATGTGTAATGTGAGAGACACAGCCCTCTCTTTCAGAGGAAAGGGGCTTGAGGTG 1978  
QY 2203 CCAGCTGTGCTCTGGGTTAGGGTTGGGGTTCGGGCTTCAGGGCCAGGAAGCAGGT 2262  
DB 1979 CCAGCTGTGCTCTGGGTTAGGGTTGGGGTTCGGGCTTCAGGGCCAGGAAGCAGGT 2038  
QY 2263 TCCCTCTCTGTGCTGTCTTGTGCAAGTCTTTAGAGGAAATAAAAAGGAAAGTGAG 2317  
DB 2039 TCCCTCTCTGTGCTGTCTTGTGCAAGTCTTTAGAGGAAATAAAAAGGAAAGTGAG 2093

RESULT 7  
ACC72138  
ID ACC72138 standard; DNA; 2105 BP.  
XX

AC ACC72138;  
 XX 07-JUL-2003 (first entry)  
 XX Human NOV33a coding sequence.  
 XX Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipidemic; gene therapy; metabolic disorder;  
 KW diabetes; obesity; infection; cachexia; cancer; dyslipidaemia;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX WO2003029423-A2.  
 XX  
 XX 10-APR-2003.  
 XX  
 XX 02-OCT-2002; 2002WO-US031358.  
 XX  
 XX 02-OCT-2001; 2001US-0326483P.  
 XX 05-OCT-2001; 2001US-0327342P.  
 XX 09-OCT-2001; 2001US-0327917P.  
 XX 09-OCT-2001; 2001US-0328029P.  
 XX 09-OCT-2001; 2001US-0328044P.  
 XX 09-OCT-2001; 2001US-0328056P.  
 XX 12-OCT-2001; 2001US-0328849P.  
 XX 15-OCT-2001; 2001US-0329414P.  
 XX 17-OCT-2001; 2001US-0330142P.  
 XX 22-OCT-2001; 2001US-0341058P.  
 XX 24-OCT-2001; 2001US-0339266P.  
 XX 24-OCT-2001; 2001US-0343629P.  
 XX 29-OCT-2001; 2001US-0349575P.  
 XX 01-NOV-2001; 2001US-0346357P.  
 XX 12-APR-2002; 2002US-0371972P.  
 XX 12-APR-2002; 2002US-0371980P.  
 XX 17-APR-2002; 2002US-0373261P.  
 XX 19-APR-2002; 2002US-0373805P.  
 XX 23-APR-2002; 2002US-0374738P.  
 XX 16-MAY-2002; 2002US-0381101P.  
 XX 17-MAY-2002; 2002US-0381635P.  
 XX 29-MAY-2002; 2002US-0383830P.  
 XX 01-OCT-2002; 2002US-00262839.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 XX Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;  
 PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
 PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;  
 PI Rothenberg ME, Shimkets RA, Smithson G, Spytek KA, Taupier RJ;  
 PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;  
 XX  
 XX WPI; 2003-381625/36.  
 XX P-PSDB; ABR58426.  
 XX  
 XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
 PT dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 XX Claim 20; Page 209-210; 487pp; English.  
 XX  
 XX The present invention relates to novel human NOV proteins and their  
 CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
 CC proteins are useful in manufacturing a medicament for treating a syndrome  
 CC associated with a human disease. The NOV proteins and coding sequences  
 CC may be used to diagnose, treat or prevent metabolic disorders such as  
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
 CC disorders, haematopoietic disorders and various dyslipidaemias  
 XX  
 XX Sequence 2105 BP; 425 A; 629 C; 556 G; 495 T; 0 U; 0 Other;

Query Match		82.3%;	Score 1914;	DB 7;	Length 2105;
Best Local Similarity		93.4%;	Pred. No. 0;		
Matches 2088;		Conservative	0;	Mismatches	10;
				Indels	138;
				Gaps	3;
QY	31	AGGTGCAAGCCTGTGTCGCCCCGAGCCCTGCGAGCTCGGGCCAGCATGCCGCCACGCT	90		
DB	1	AGGTGCAAGCCTGTGTCGCCCCGAGCCCTGCGAGCTCGGGCCAGCATGCCGCCACGCT	60		
QY	91	GCACAGGCGTACCGAGGCGCTGTGGATGCGCTGCAAGGCTGTGCTGGAGAACCTCTT	150		
DB	61	GCACAGGCGTACCGAGGCGCTGTGGATGCGCTGCAAGGCTGTGCTGGAGAACCTCTT	120		
QY	151	CTTCTCTGCTGATCTCTGGGCTGGGCTCCCTGTTGATCATCTTGAAGAACGAGGGCTT	210		
DB	121	CTTCTCTGCTGATCTCTGGGCTGGGCTCCCTGTTGATCATCTTGAAGAACGAGGGCTT	180		
QY	211	CTATTCCAGCAGCTGCCAGCTGAGAGCAGCAACCAACACCCAGGATGAGCAGCGCAG	270		
DB	181	CTATTCCAGCAGCTGCCAGCTGAGAGCAGCAACCAACACCCAGGATGAGCAGCGCAG	240		
QY	271	GTGGCCAGGCTGTGACCAGCAGGATGCTCAACCTGGGCTTCACCATTTGGTTCCCTT	330		
DB	241	GTGGC-----	245		
QY	331	CGTGTCTCAGCGCCACACCTGCCACTGGGGATCTCTATGGACCGCTTTGGCCCCCGAC	390		
DB	246	-----	245		
QY	391	CGTGGGCTGGTTGGCAGTGCCTGTTCACTCGCTGCTGCAACCTCATGGCCCTGCGCTC	450		
DB	246	-----CTTGTCTTCACTGCGTCTGCAACCTCATGGCCCTGCGCTC	285		
QY	451	CGGGAGCTGGAAGCTCTGCTCCCTGATATTCCTGGGCTGCTCCCTGAATGGCTTTCG	510		
DB	286	CGGGAGCTGGAAGCTCTGCTCCCTGATATTCCTGGGCTGCTCCCTGAATGGCTTTCG	345		
QY	511	TGGCATCTGCTTAACGCTTCACTTCACTCACGCTGCCCAACATGTTTGGAACTCGGCTC	570		
DB	346	TGGCATCTGCTTAACGCTTCACTTCACTCACGCTGCCCAACATGTTTGGAACTCGGCTC	405		
QY	571	CACGTTAATGGCCCTCATGATGGCTCTTAAACGCTTCTGTCATTAAGTTCCAGGAAT	630		
DB	406	CACGTTAATGGCCCTCATGATGGCTCTTAAACGCTTCTGTCATTAAGTTCCAGGAAT	465		
QY	631	CAAGCTGATCTACGATGCGGCTTGGCTTCTGCTGTCATCATGTTCACTGGCTGGCCT	690		
DB	466	CAAGCTGATCTACGATGCGGCTTGGCTTCTGCTGTCATCATGTTCACTGGCTGGCCT	525		
QY	691	GGCTTGCCTTATCTTTCTGAACTGCACACCTCAACTGGGCCCATCGAAGCCTTTCTGCCCC	750		
DB	526	GGCTTGCCTTATCTTTCTGAACTGCACACCTCAACTGGGCCCATCGAAGCCTTTCTGCCCC	585		
QY	751	TGAGGAAGTCAATTACAGGAAGATCAAGCTGATGGGCTGGCCCTGGACCAAGGT	810		
DB	586	TGAGGAAGTCAATTACAGGAAGATCAAGCTGATGGGCTGGCCCTGGACCAAGGT	645		
QY	811	GACAGGTGACCTTCTTACACCCATGTGACACCATGGGCCAGAGGCTCAGCAGAGGC	870		
DB	646	GACAGGTGACCTTCTTACACCCATGTGACACCATGGGCCAGAGGCTCAGCAGAGGC	705		
QY	871	CCCCAGCTGGAGGACGGTTGGATGCTTCTATGTACCCAGGATGTTGGGGGACCTTC	930		
DB	706	CCCCAGCTGGAGGACGGTTGGATGCTTCTATGTACCCAGGATGTTGGGGGACCTTC	765		
QY	931	AGAAAACCTTCTGTAGAGGTCTGTCCTTACGCAAGAGCCTCTGCTCCCCACCTTCTCT	990		
DB	766	AGAAAACCTTCTGTAGAGGTCTGTCCTTACGCAAGAGCCTCTGCTCCCCACCTTCTCT	825		
QY	991	GTGGAGCTCTCTCACCATGGGCGATGACCCAGCTGGGATCATCTTCTACATGGCTGCTGT	1050		
DB	826	GTGGAGCTCTCTCACCATGGGCGATGACCCAGCTGGGATCATCTTCTACATGGCTGCTGT	885		

QY 1051 GAAACAGATGCTGGAGTACCTTGTGACTGGTGGCCAGGAGATGAGACAAATGAACAGCA 1110  
 DB 886 GAAACAGATGCTGGAGTACCTTGTGACTGGTGGCCAGGAGATGAGACAAATGAACAGCA 945  
 QY 1111 ACAAAAGTGGCAGAGACAGTGGGTCTTATCTCTCCGTCTTTCGGGGCATGACAGCTGTT 1170  
 DB 946 ACAAAAGTGGCAGAGACAGTGGGTCTTATCTCTCCGTCTTTCGGGGCATGACAGCTGTT 1005  
 QY 1171 GTGCCCTTCTACCTGCCCTCATTTGGCTATCATGACTGGCGGATCAAGGACTGGT 1230  
 DB 1006 GTGCCCTTCTACCTGCCCTCATTTGGCTATCATGACTGGCGGATCAAGGACTGGT 1065  
 QY 1231 GGACGCCCACTCAGGGGCACTGTCTCGGAGATGCCAGGGACGGGTGCTACCAATC 1290  
 DB 1066 GGACGCCCACTCAGGGGCACTGTCTCGGAGATGCCAGGGACGGGTGCTACCAATC 1125  
 QY 1291 CATCAGACACGCTACTGCAAGATCCAAAGCTCAACCAATGCTCATGCTTCCACCTT 1350  
 DB 1126 CATCAGACACGCTACTGCAAGATCCAAAGCTCAACCAATGCTCATGCTTCCACCTT 1185  
 QY 1411 GTTGTGACCTTGTCTGCACACCAATGTTGCGAGTCTTTCACACTCAGCCTGTGGAG 1470  
 DB 1246 GTTGTGACCTTGTCTGCACACCAATGTTGCGAGTCTTTCACACTCAGCCTGTGGAG 1305  
 QY 1471 TCTTATGCTGAGTGTTCCTCCATCAACCACTTTGGAGCGTGCAGGCTGCACTCCCT 1530  
 DB 1306 TCTTATGCTGAGTGTTCCTCCATCAACCACTTTGGAGCGTGCAGGCTGCACTCCCT 1365  
 QY 1531 CATCAGTCTGTGTTGGCTGCTTTCAGAGCCTTTCATGCGGATGTTGGAGCCCT 1590  
 DB 1366 CATCAGTCTGTGTTGGCTGCTTTCAGAGCCTTTCATGCGGATGTTGGAGCCCT 1425  
 QY 1591 GAAAGGAGAGCCCTTCTGGGTGAATCTGGGCTCTGCTATTTCTCACTCTCGGATTCCT 1650  
 DB 1426 GAAAGGAGAGCCCTTCTGGGTGAATCTGGGCTCTGCTATTTCTCACTCTCGGATTCCT 1485  
 QY 1651 GTTGTCTTCTACCTTCTTATTAACGTCGCGCTCCAGCAGGAGTACGCGCCCAATGG 1710  
 DB 1486 GTTGTCTTCTACCTTCTTATTAACGTCGCGCTCCAGCAGGAGTACGCGCCCAATGG 1545  
 QY 1711 GATGGGCCACCTGAGAGTCTTAGCGGCTCTGAGTGACCGCATAGCTTCTCAGACCAA 1770  
 DB 1546 GATGGGCCACCTGAGAGTCTTAGCGGCTCTGAGTGACCGCATAGCTTCTCAGACCAA 1605  
 QY 1771 GGGACCTGGATGACAGGCAATCAAGGCTCTGAGCAACCAAGGAGTGGCCCATATGGCTT 1830  
 DB 1606 GGGACCTGGATGACAGGCAATCAAGGCTCTGAGCAACCAAGGAGTGGCCCATATGGCTT 1665  
 QY 1831 TTCTACCTGTAAATGACATAGAGCCTATGGCGGTAGATTTTAAATACCAAGAGAAGTT 1890  
 DB 1666 TTCTACCTGTAAATGACATAGAGCCTATGGCGGTAGATTTTAAATACCAAGAGAAGTT 1725  
 QY 1891 CTATTTTGTAAAGTGCAGAAAGGAGG-AAAAAACCCTTCAAAAAGCCCTTAACT 1949  
 DB 1726 CTATTTTGTAAAGTGCAGAAAGGAGGAGGAAAAAAACCTTCAAAAAGCCCTTAACT 1785  
 QY 1950 CAACGCTCATTTGACTGAAGACAGTCCCTATCTAGAGGGGTGAGCTTTCTTCTCTCTT 2009  
 DB 1786 CAACGCTCATTTGACTGAAGACAGTCCCTATCTAGAGGGGTGAGCTTTCTTCTCTCTT 1845  
 QY 2010 GGGTGGAGGAGACAGGGTGCCTTATCTCTTCTAGCGGTCTGCTCTCTGTTACCTC 2069  
 DB 1846 GGGTGGAGGAGACAGGGTGCCTTATCTCTTCTAGCGGTCTGCTCTCTGTTACCTC 1905  
 QY 2070 TTGGGGGATCGGCAAAACAGGCTACCCCTGAGGTCCCATGTCATGAGTGTGCAACA 2129  
 DB 1906 TTGGGGGATCGGCAAAACAGGCTACCCCTGAGGTCCCATGTCATGAGTGTGCAACA--C 1963  
 QY 2130 TGCATGTGTGTGTGTGTGAATGTGAGAAAAACACAGCCCTCTTTTTCAGAGGAAA 2189

DB 1964 ATGCATGTGTGTGTGTGAATGTGAGAGACACAGCCCTCTCTTTCAGAGGAAA 2023  
 QY 2190 GGGGCTCAGTGCACAGCTGTCTCTGGTTAGGGTTGGGGTTCGGCCCTTCCAGGCG 2249  
 DB 2024 GGGGCTCAGTGCACAGCTGTCTCTGGTTAGGGTTGGGGTTCGGCCCTTCCAGGCG 2083  
 QY 2250 CAGGAAGCAGGTTCC 2265  
 DB 2084 CAGGAGTCAAGTTCC 2099  
 RESULT 8  
 ID AAV49591 standard; cDNA to mRNA; 1677 BP.  
 AC AAV49591;  
 XX 21-OCT-1998 (first entry)  
 XX Human liver cell clone HP10302 cDNA #2.  
 XX Transmembrane domain; human; nutrition; cytokine; cell proliferation;  
 KW differentiation; immune system; stimulator; suppressor; regulator;  
 KW hematopoiesis; activin; inhibitor; chemotactic; chemokinetic; receptor;  
 KW haemostatic; thrombolytic; ligand; anti-inflammatory; tumour; ds.  
 XX Homo sapiens.  
 OS  
 XX WO9821328-A2.  
 PN  
 XX 22-MAY-1998.  
 PD  
 XX 07-NOV-1997; 97WO-JP004056.  
 PF  
 XX 13-NOV-1996; 96JP-00301429.  
 PR  
 XX (SAGA) SAGAMI CHEM RES CENTRE.  
 PA (PROT-) PROTEGENE INC.  
 XX  
 PI Kato S, Sekine S, Yamaguchi T, Kobayashi M;  
 XX WPI; 1998-297932/26.  
 DR P-PSDB; AAW64554.  
 XX  
 PT Human protein having transmembrane domain - useful for, e.g. research and  
 PT nutrition.  
 XX  
 PS Claim 3; Page 130; 205pp; English.  
 XX  
 CC AAV49550-V49599 are cDNA sequences which encode human proteins containing  
 CC a transmembrane domain. These proteins can be used for, e.g. research and  
 CC nutrition, and may have cytokine and cell proliferation/differentiation,  
 CC immune stimulating/suppressing, haematopoiesis regulating, tissue growth,  
 CC activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic,  
 CC receptor/ligand, anti-inflammatory or tumour inhibition activity  
 XX  
 SQ Sequence 1677 BP; 310 A; 534 C; 436 G; 397 T; 0 U; 0 Other;  
 Query Match 72.1%; Score 1677; DB 2; Length 1677;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 77 ATGGCCCCCAGCTGCAACAGGCGTACCGGAGGCGCTGGTGGTGGCTGCACGCTGTG 136  
 DB 1 ATGGCCCCCAGCTGCAACAGGCGTACCGGAGGCGCTGGTGGTGGCTGCACGCTGTG 60  
 QY 137 CTGGAGAACCTCTTCTTCTCTGCTACTCTCTGGGCTGGGCTCCCTGTTGATCATTTCTG 196  
 DB 61 CTGGAGAACCTCTTCTTCTCTGCTACTCTCTGGGCTGGGCTCCCTGTTGATCATTTCTG 120  
 QY 197 AAGAACAGGCGCTTCTATTCCAGCAGCTGCCAGCTGAGAGCAGCACCAACCCAG 256

Db 121 AAGAACGAGGGCTTCTATTCCAGCAGCTGCCAGCTGAGAGCAGCAGCAACCAACCCAG 180  
 QY 257 GATGACAGCGCAGGTGGCCAGGCTGTGACACGAGCAGAGATGCTCAACCTGGGCTTC 316  
 Db 181 GATGACAGCGCAGGTGGCCAGGCTGTGACACGAGCAGAGATGCTCAACCTGGGCTTC 240  
 QY 317 ACCATTGGTTCCTTCTGCTGACGCGCACACCCCTGCCACTCGGGATCTCATGGACCGC 376  
 Db 241 ACCATTGGTTCCTTCTGCTGACGCGCACACCCCTGCCACTCGGGATCTCATGGACCGC 300  
 QY 377 TTTGGCCCCGAGCCCTGGCTGGTGGCAGTGGCTTCACTGGCTCCTGCACCCCTC 436  
 Db 301 TTTGGCCCCGAGCCCTGGCTGGTGGCAGTGGCTTCACTGGCTCCTGCACCCCTC 360  
 QY 437 ATGGCCCTGGCTCCTCCCGGACGTGGAGCTCTGTCTCCGTTGATATTTCTGGCGCTGCC 496  
 Db 361 ATGGCCCTGGCTCCTCCCGGACGTGGAGCTCTGTCTCCGTTGATATTTCTGGCGCTGCC 420  
 QY 497 CTGAATGGCTTTGGTGGCATCTGCCCTAACGTTCACTTCACTCACTGCTGCCAACAATGTTT 556  
 Db 421 CTGAATGGCTTTGGTGGCATCTGCCCTAACGTTCACTTCACTCACTGCTGCCAACAATGTTT 480  
 QY 557 GGGAACTCGCTCCAGCTTAATGGCCCTCATGATTTGGCTCTTACGCCCTCTTCTGCCATT 616  
 Db 481 GGGAACTCGCTCCAGCTTAATGGCCCTCATGATTTGGCTCTTACGCCCTCTTCTGCCATT 540  
 QY 617 ACCTTCCAGGAATCAAGCTGATCTACGATGCCGCTGGCCCTTCTGTGTCATCATGTTTC 676  
 Db 541 ACCTTCCAGGAATCAAGCTGATCTACGATGCCGCTGGCCCTTCTGTGTCATCATGTTTC 600  
 QY 677 ACCTGCTGTGGCTGGCTGCCCTTATCTTTTCTGAACCTGCACCTCAACTGGGCCATCGAA 736  
 Db 601 ACCTGCTGTGGCTGGCTGCCCTTATCTTTTCTGAACCTGCACCTCAACTGGGCCATCGAA 660  
 QY 737 GCTTTCTCCCTCCCTCAGGAGTCAATTACAGAGAGATCAAGCTGAGTGGCTGGCC 796  
 Db 661 GCTTTCTCCCTCCCTCAGGAGTCAATTACAGAGAGATCAAGCTGAGTGGCTGGCC 720  
 QY 797 CTGGACCAAGGTGACAGTGAACCTCTCTGAGAGTCTGTCCCTTAGCGAAGACCTCTGC 856  
 Db 721 CTGGACCAAGGTGACAGTGAACCTCTCTGAGAGTCTGTCCCTTAGCGAAGACCTCTGC 780  
 QY 857 CTGAGCAGAGGCCCCAGCTGGAGGCGGTTCCGATGCTTCACTGTCACCCAGGAT 916  
 Db 781 CTGAGCAGAGGCCCCAGCTGGAGGCGGTTCCGATGCTTCACTGTCACCCAGGAT 840  
 QY 917 GTTCGGGGACCTCAGAAAACCTTCTGAGAGTCTGTCCCTTAGCGAAGACCTCTGC 976  
 Db 841 GTTCGGGGACCTCAGAAAACCTTCTGAGAGTCTGTCCCTTAGCGAAGACCTCTGC 900  
 QY 977 TCCCCCAGCTTCTGTTGGAGCTCTCTCAGCGGATGACCCAGCTGGGATCATCTTC 1036  
 Db 901 TCCCCCAGCTTCTGTTGGAGCTCTCTCAGCGGATGACCCAGCTGGGATCATCTTC 960  
 QY 1037 TACATGGCTGTGTGAACAAGATGCTGGAGTACCTTTGACCTGGTGGCCAGGAGCATGAG 1096  
 Db 961 TACATGGCTGTGTGAACAAGATGCTGGAGTACCTTTGACCTGGTGGCCAGGAGCATGAG 1020  
 QY 1097 ACAATGAACAGCAACAAAGGTGGCAGAGACAGTTGGGTTCTACTCTCCCTCTTCGGG 1156  
 Db 1021 ACAATGAACAGCAACAAAGGTGGCAGAGACAGTTGGGTTCTACTCTCCCTCTTCGGG 1080  
 QY 1157 GCCATGACGCTGTTGGCTTCTCAGCTGCCCTCATTTGGCTTACATCATGACTGGCGG 1216  
 Db 1081 GCCATGACGCTGTTGGCTTCTCAGCTGCCCTCATTTGGCTTACATCATGACTGGCGG 1140  
 QY 1217 ATCAAGGACTGGTGGACGCCCAACTCAGGGCACTGTCTCTCGAGATGCCAGGACGGG 1276  
 Db 1141 ATCAAGGACTGGTGGACGCCCAACTCAGGGCACTGTCTCTCGAGATGCCAGGACGGG 1200  
 QY 1277 GTTGCTACCAATTCATCAGACGACTACTGAGAGTCAAAAGCTCAACCAATGCCATC 1336  
 Db 1201 GTTGCTACCAATTCATCAGACGACTACTGAGAGTCAAAAGCTCAACCAATGCCATC 1260

QY 1337 AGTGCCTTCAACCTGACCAACCTGCTGTGTGGGTTTTGGCATCACCTGTCTCATCAAC 1396  
 Db 1261 AGTGCCTTCAACCTGACCAACCTGCTGTGTGGGTTTTGGCATCACCTGTCTCATCAAC 1320  
 QY 1397 AACTTACACCTCCAGTTTGTGACCTTTGTCTGTCGACACCAATTTGTCGAGGTTTCTCCAC 1456  
 Db 1321 AACTTACACCTCCAGTTTGTGACCTTTGTCTGTCGACACCAATTTGTCGAGGTTTCTCCAC 1380  
 QY 1457 TCAGCTGTGGGAGTCTCATGCTGCAGTGTTCGCCCTTGTCTTACAGCAGCCACTTTTTCATGGCG 1516  
 Db 1381 TCAGCTGTGGGAGTCTCATGCTGCAGTGTTCGCCCTTGTCTTACAGCAGCCACTTTTTCATGGCG 1440  
 QY 1517 GGCCTGCACTCCCTCATCAGTGTGTGTGGCTTGTCTTACAGCAGCCACTTTTTCATGGCG 1576  
 Db 1441 GGCCTGCACTCCCTCATCAGTGTGTGTGGCTTGTCTTACAGCAGCCACTTTTTCATGGCG 1500  
 QY 1577 ATGTGGGAGCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGGCCCTCTGCTATTCTCA 1636  
 Db 1501 ATGTGGGAGCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGGCCCTCTGCTATTCTCA 1560  
 QY 1637 CTCCTGGGATTCCTTGTCTTACCTTCTTATACCGTGGCCGCTCCAGCAGGAG 1696  
 Db 1561 CTCCTGGGATTCCTTGTCTTACCTTCTTATACCGTGGCCGCTCCAGCAGGAG 1620  
 QY 1697 TACGCGCGCAATGGGATGGGATGGCCCACTGAAGGTGCTTAGCGGCTCTGAGGTGACCGCA 1753  
 Db 1621 TACGCGCGCAATGGGATGGGATGGCCCACTGAAGGTGCTTAGCGGCTCTGAGGTGACCGCA 1677

RESULT 9

AAZ50443  
ID AAZ50443 standard; cDNA; 3442 bp.

AC AAZ50443;

XX 18-MAY-2000 (first entry)

XX Human PB39 variant cDNA, a gene dysregulated in prostate cancer.

XX PB39; human; prostate cancer; PC; chromosome 11p11.1-11.2; variant;  
 prostate epithelium; splicing mechanism; early diagnosis; progression;  
 precancerous cell; aggressive prostate carcinoma; metastatic potential;  
 cancer; non-neoplastic prostate disease; expressed sequence tag; EST; ss.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH 1760..3442  
 CDS /\*tag= a

FT /product= "Human PB39 variant protein"  
 FT /note= "Derived from alternative splicing of mRNA"

XX WO200005376-A1.

XX 03-FEB-2000.

XX 23-JUL-1999; 99WO-US016831.

XX 24-JUL-1998; 98US-0094137P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Chuaqui RF, Cole KA, Liotta LA;

XX WPI; 2000-182700/16.

XX P-PSDB; AAY44898.

XX Novel gene which is dysregulated in prostate cancer useful for diagnosing  
 cancer.

XX Claim 1; Page 41-44; 51pp; English.

XX



The present sequence is the human PB39 variant cDNA, a gene that is dysregulated in prostate cancer. It is isolated from human pancreas cDNA library and has homology to an expressed sequence tag (EST) AAR00504. PB39 variant results from an alternative RNA splicing mechanism, yielding a larger transcript (5KB). It is located on the chromosome 11p11.1-11.2. Abnormally high concentrations of PB39 are found in prostate tissue derived from prostate cancer (PC) epithelium. PB39 sequence is useful for detection of precancerous or cancer cells in the prostate. PB39 is useful for early diagnosis of the progression of prostate cancer, especially in aggressive prostate carcinoma. It can also distinguish PC from other non-neoplastic prostate disease. The diagnostic method is selective and specific for various types of PC and also facilitates identifying prostate cancer of differing aggressiveness and metastatic potential.

Sequence 3442 BP; 636 A; 1071 C; 933 G; 802 T; 0 U; 0 Other;

Query Match 69.3%; Score 1613; DB 3; Length 3442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1613; Conservative 0; Mismatches 0; Indels 0

Qy	1	CGGGGCTGGAGGGGGGCAAGCGGGTTCGAGAGTGCAAAGCTGTGTGTCGCCGAGACCCCTGC	60
Db	1	CGGGGCTGGAGGGGGGCAAGCGGGTTCGAGAGTGCAAAGCTGTGTGTCGCCGAGACCCCTGC	60
Qy	61	GGAGCTCGGGGCGACGATGGCGCCCAAGCGCTGCAACAGCGCTACCGAGGGCGCTGTGGAT	120
Db	61	GGAGCTCGGGGCGACGATGGCGCCCAAGCGCTGCAACAGCGCTACCGAGGGCGCTGTGGAT	120
Qy	121	GGCTGCACGGCTGTGCTGGAGAACCTCTTCTCTCTGTATCTCTGGGCTGGGGCTC	180
Db	121	GGCTGCACGGCTGTGCTGGAGAACCTCTTCTCTCTGTATCTCTGGGCTGGGGCTC	180
Qy	181	CTGTGTGATCATCTGAAGAACGAGGGCTCTATTTCAGCAGCTGCCAGCTGAGAGCAG	240
Db	181	CTGTGTGATCATCTGAAGAACGAGGGCTCTATTTCAGCAGCTGCCAGCTGAGAGCAG	240
Qy	241	CACCAACACACCCAGGATGAGCAGCGCAGGTGGCAGGCTGTGACCAAGCAGGACGAGAT	300
Db	241	CACCAACACACCCAGGATGAGCAGCGCAGGTGGCAGGCTGTGACCAAGCAGGACGAGAT	300
Qy	301	GCTCAACCTGGGCTTCAACATTTGTTCTCTGTCTCAGCGCCACCAACCTGCCACTGGG	360
Db	301	GCTCAACCTGGGCTTCAACATTTGTTCTCTGTCTCAGCGCCACCAACCTGCCACTGGG	360
Qy	361	GATCTCATGAGACCGTTTGGCCCCGAGCCGCTGGGCTGGTGGCAGTGCCTGCTTCAC	420
Db	361	GATCTCATGAGACCGTTTGGCCCCGAGCCGCTGGGCTGGTGGCAGTGCCTGCTTCAC	420
Qy	421	TGGCTCTCACCCCTCATGGCCCTGGCCCTCCGGGAGCTGGAAGCTCTGTCTCCGTTGAT	480
Db	421	TGGCTCTCACCCCTCATGGCCCTGGCCCTCCGGGAGCTGGAAGCTCTGTCTCCGTTGAT	480
Qy	481	ATTCTCGGCGCTGCTCCCTGAATGGCTTTGTTGGGATCTGCTCAACGTTCACTTCACTCAC	540
Db	481	ATTCTCGGCGCTGCTCCCTGAATGGCTTTGTTGGGATCTGCTCAACGTTCACTTCACTCAC	540
Qy	541	GCTGCCCAACATGTTTGGGACCTTGGCTCCACGTTAATGGCCCTCATGATGGCTCTTA	600
Db	541	GCTGCCCAACATGTTTGGGACCTTGGCTCCACGTTAATGGCCCTCATGATGGCTCTTA	600
Qy	601	CGCCTCTTCGCAATTAAGTTTCCAGGAATCAAGCTGATCTACGATCCGGTGTGGCCTT	660
Db	601	CGCCTCTTCGCAATTAAGTTTCCAGGAATCAAGCTGATCTACGATCCGGTGTGGCCTT	660
Qy	661	CGTGTGCATCATGTTCACTGGTCTGGGCTGGCCTGCTTATCTTTCTGAACCTGCACCT	720
Db	661	CGTGTGCATCATGTTCACTGGTCTGGGCTGGCCTGCTTATCTTTCTGAACCTGCACCT	720
Qy	721	CAACTGGCCCCATCGAAGCCTTTCGCCCCGAGGAGTCAATTTACACGAAGAGATCAA	780
Db	721	CAACTGGCCCCATCGAAGCCTTTCGCCCCGAGGAGTCAATTTACACGAAGAGATCAA	780
Qy	781	GCTGAGTGGGCTGGCCCTGGACCAAGGTCAGAGGTGACCTCTTCTTACACCCATGTGAC	840



PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US0001339.  
XX PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 18-AUG-2000; 2000US-0225759P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227700P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234397P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246529P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250191P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251985P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465570/50.  
DR

XX Isolated nucleic acid molecule encoding a reproductive system antigen is  
PT used in preventing, treating or ameliorating a medical condition.  
XX  
XX Disclosure; SEQ ID NO 8087; 1297pp + Sequence Listing; English.  
PS  
XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention  
XX  
SQ Sequence 1737 BP; 443 A; 431 C; 458 G; 405 T; 0 U; 0 Other;  
  
Query Match 29.4%; Score 684.2; DB 4; Length 1737;  
Best Local Similarity 97.9%; Pred. No. 1.1e-166;  
Matches 704; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
  
QY 1602 CCTTCTGGGTGAATCTGGGCTCTCTGCTATTCTCACTCTGGGATTCCTGTCCTTCT 1661  
DB 1189 CCTTAAGGTGAATCTGGGCTCTCTGCTATTCTCACTCTGGGATTCCTGTCCTTCT 1130  
QY 1662 ACCTCTTCTATACCGTCCCGCTCCAGCAGAGTACGCCCAATGGATGGGCCAC 1721  
DB 1129 ACCTCTTCTATACCGTCCCGCTCCAGCAGAGTACGCCCAATGGATGGGCCAC 1070  
QY 1722 TGAAGTCTTAGCGCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1781  
DB 1069 TGAAGTCTTAGCGCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1010  
QY 1782 GACGCAATCAAGCCTGAGCAACCAAAAGAGTGCCCATATGGCTTTTCTACTCTGA 1841  
DB 1009 GACGCAATCAAGCCTGAGCAACCAAAAGAGTGCCCATATGGCTTTTCTACTCTGA 950  
QY 1842 ACATGCACATAGAGCAGTCCGCTAGATTATATAATACCAAGAGATTCATTTTGTG 1901  
DB 949 ACATGCACATAGAGCAGTCCGCTAGATTATATAATACCAAGAGATTCATTTTGTG 890  
QY 1902 AAGACTGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1961  
DB 889 AAGACTGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 830  
QY 1962 GACTGAAGACAGTCCCTATCTCTAGAGGGTTCAGCTTTCTCTCTCTCTCTCTCTCT 2021  
DB 829 GACTGAAGACAGTCCCTATCTCTAGAGGGTTCAGCTTTCTCTCTCTCTCTCTCTCT 770  
QY 2022 ACCAGGTGCTCTTATCT 2081  
DB 769 ACCAGGTGCTCTTATCT 710  
QY 2082 GCAAAAGGCTACCCCTGAGGTCCCATGTCATGAGTGTGCACATGCAATGTGTCT 2141  
DB 709 GCAAAAGGCTACCCCTGAGGTCCCATGTCATGAGTGTGCACA - CATGCAATGTGTCT 652  
QY 2142 GTGTATGTGTGAATGTGAAGAAACACAGCCCTCTCTCTCAGAGGAGGAGGAGGAGG 2201  
DB 651 GTGTATGTGTGAATGTGAAGAGACACAGCCCTCTCTCTCAGAGGAGGAGGAGGAGG 592  
QY 2202 GCCAGTGTGTCTCTGGTGTAGGGTGTGGGGTGTGGGGTGTGGGGTGTGGGGTGTGGGG 2261  
DB 591 GCCAGTGTGTCTCTGGTGTAGGGTGTGGGGTGTGGGGTGTGGGGTGTGGGGTGTGGGG 532  
QY 2262 TTCCCTCTCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 2320  
DB 531 TTCCCTCTCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 473

RESULT 11  
ABL98261/c  
ID ABL98261 standard; DNA; 1737 BP.  
XX  
AC ABL98261;  
XX

DT 21-JUN-2002 (first entry)  
DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2913.  
XX  
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
XX gastrointestinal disease; infection; cytostatic; gene; da.  
OS Homo sapiens.  
XX  
FN WO200155317-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001329.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214896P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216680P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 08-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.

```
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237003P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239353P.
PR 13-OCT-2000; 2000US-0239357P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0246177P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful
XX for preventing, diagnosing and/or treating testicular cancer.
XX
XX Disclosure; SEQ ID NO 2913; 766pp; English.
XX
XX The present invention provides the protein and coding sequences of 973
XX human testicular antigens, and fragments of their genomic sequences. The
XX sequences can be used in the treatment of cardiovascular, urinary system,
XX reproductive system, immune, respiratory, neurological and
XX gastrointestinal disorders, infections, and particularly cancer,
XX especially testicular cancers. The present sequence is a DNA encoding a
XX protein fragment of the invention
XX
XX Sequence 1737 BP; 443 A; 431 C; 458 G; 405 T; 0 U; 0 Other;
XX

Query Match 29.4%; Score 684.2; DB 4; Length 1737;
Best Local Similarity 97.9%; Pred. No. 1.1e-166;
Matches 704; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1602 CTTCTGGGTGAATCTGGGCTCTCTGCTATTCTCACTCTGGGATTCCTGTCCTTCT 1661
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1662 ACCTCTTCTATTACCGTGCCGGCTCCAGCAGGAGTACGCCCAATGGGATGGGCCAC 1721
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1129 ACCTCTTCTATTACCGTGCCGGCTCCAGCAGGAGTACGCCCAATGGGATGGGCCAC 1070
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1722 TGAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1781
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1069 TGAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1010
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1782 GACAGGCAATCAAGGCTGAGCAACCAAGAGTGCCCATATGGCTTTCTACCTGTA 1841
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1009 GACAGGCAATCAAGGCTGAGCAACCAAGAGTGCCCATATGGCTTTCTACCTGTA 950
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1842 ACATGCACATAGAGCCATGGCGTAGATTTATAAATACCAAGAGAGTTCTATTTTGTGA 1901
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 949 ACATGCACATAGAGCCATGGCGTAGATTTATAAATACCAAGAGAGTTCTATTTTGTGA 890
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1902 AAGACTGCAAAAGGAGGAAAAAACCTTCAAAACGCCCTTAAGTCAACGCTCATT 1961
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 889 AAGACTGCAAAAGGAGGAAAAAACCTTCAAAACGCCCTTAAGTCAACGCTCATT 830
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1962 GACTGAACACAGTCCCTATCTTAGAGGGGTGAGCTTTCTCTCTGGGTGGAGGAG 2021
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 829 GACTGAACACAGTCCCTATCTTAGAGGGGTGAGCTTTCTCTCTGGGTGGAGGAG 770
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2022 ACCAGGTGCTCTTATCTCTTAGCGGTCTGCTCTCTGCTCTCTGCTCTCTGCTCTGCT 2081
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 769 ACCAGGTGCTCTTATCTCTTAGCGGTCTGCTCTCTGCTCTCTGCTCTCTGCTCTGCT 710
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2082 GCAAAACAGGCTACCCCTGAGTCCCATGTGCCATGAGTGTGCACATGCAATGTCT 2141
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 709 GCAAAACAGGCTACCCCTGAGTCCCATGTGCCATGAGTGTGCACATGCAATGTCT 652
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2142 GTGTATGTGTAATGTGAGAAAAACACAGCCCTCTTTTTCAGAAAGGAGGCTGAGGT 2201
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 651 GTGTATGTGTAATGTGAGAGACACAGCCCTCTTTTTCAGAAAGGAGGCTGAGGT 592
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Qy 2202 GCCAGCTGTCTCTGGTGTAGGGTTGGGGTTGGGGTTCGCCCTTCCAGGGCCAGGAAGGCAGG 2261  
Db 591 GCCAGCTGTCTCTGGTGTAGGGTTGGGGTTGGGGTTCGCCCTTCCAGGGCCAGGAAGGCAGG 532  
Qy 2262 TTCCCTCTCTGGTGTCTGTCTGTGCTTGAAGTCTTAGAGGAAATAAAAGGAAAGTGAGAAA 2320  
Db 531 TTCCCTCTCTGGTGTCTGTCTGTGCTTGAAGTCTTAGAGGAAATAAAAGGAAAGTGAGAGA 473

RESULT 12

AAAL05398/c  
ID AAL05398 standard; DNA; 3160 BP.

XX AC AAL05398;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen DNA SEQ ID NO: 8086.

XX KW Human; reproductive system related antigen; reproductive system disorder;  
XX KW cancer; gene therapy; ds.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001339.

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225211P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251388P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-465570/50.  
XX  
XX Isolated nucleic acid molecule encoding a reproductive system antigen is  
PT used in preventing, treating or ameliorating a medical condition.  
XX  
XX Disclosure; SEQ ID NO 8086; 1297pp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention  
XX  
SQ Sequence 3160 BP; 861 A; 761 C; 855 G; 683 T; 0 U; 0 Other;

Query Match 29.4%; Score 684.2; DB 4; Length 3160;  
Best Local Similarity 97.9%; Pred. No. 1.5e-166;  
Matches 704; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1602 CCTTCTGGGTGAATCTGGGCTCTCTGCTATCTCACTCTGGGATTCCTGTTGCCCTTCT 1661  
DB 1190 CCTAAAGTGAATCTGGGCTCTCTGCTATCTCACTCTGGGATTCCTGTTGCCCTTCT 1131  
QY 1662 ACCTCTTCTATTACCGTCCCGCTCCAGCAGGATACCGCCCAATGGGATGGGCCAC 1721  
DB 1130 ACCTCTTCTATTACCGTCCCGCTCCAGCAGGATACCGCCCAATGGGATGGGCCAC 1071  
QY 1722 TGAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1781  
DB 1070 TGAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1011  
QY 1782 GACAGCAATCAAGGCTCAGCAACCAAGGAGTGCCCATATGCTTTCTACCTGTA 1841  
DB 1010 GACAGCAATCAAGGCTCAGCAACCAAGGAGTGCCCATATGCTTTCTACCTGTA 951  
QY 1842 ACATGCACATAGAGCCATGGCCGTAGATTATATAATACCAAGAGAGTTCTATTTTGTGA 1901  
DB 950 ACATGCACATAGAGCCATGGCCGTAGATTATATAATACCAAGAGAGTTCTATTTTGTGA 891  
QY 1902 AGACTGCAAAAAGGAGGAAAAAAACCTTCAAAAACGCCCTTAAGTCAACGCTCCATT 1961  
DB 890 AGACTGCAAAAAGGAGGAAAAAAACCTTCAAAAACGCCCTTAAGTCAACGCTCCATT 831  
QY 1962 GACTGAAGACAGTCCCTATCTAGAGGGGTGAGCTTTCTCTCTCGGTTGGAGGAG 2021  
DB 830 GACTGAAGACAGTCCCTATCTAGAGGGGTGAGCTTTCTCTCTCGGTTGGAGGAG 771  
QY 2022 ACCAGGGTGCCTTATCTCTTCTAGCGGTCTGCTCTGCTGCTGCTTCTTGGGGGATCG 2081

Db 770 ACCAGGGTGCCTTATCTCTTCTAGCGGTCTGCTCTGCTGCTCTTGGGGGATCG 711  
QY 2082 GCAAAACAGGCTACCCCTGAGGTCCCATGTGCGCATGATGTCACACATGCAATGTGCT 2141  
Db 710 GCAAAACAGGCTACCCCTGAGGTCCCATGTGCGCATGATGTCACACATGTCATGTC 653  
QY 2142 GTGTATGTGTAATGTGAGAAAAACACAGCCCTCTCTTTCAGAGGAAAGGGGCTGAGGT 2201  
Db 652 GTGTATGTGTAATGTGAGAGAGACACAGCCCTCTCTTTCAGAGGAAAGGGGCTGAGGT 593  
QY 2202 GCCAGCTGTGCTCGGTTAGGGGTTGGGGTTCGGGCTTCAGGGCCAGGAAGGACGAG 2261  
Db 592 GCCAGCTGTGCTCGGTTAGGGGTTGGGGTTCGGGCTTCAGGGCCAGGAAGGACGAG 533  
QY 2262 TTCCCTCTCTGCTGCTGCTGCTTCAAGTCTTTAGAGGAATAAAAGGGAAGTGAGAA 2320  
Db 532 TTCCCTCTCTGCTGCTGCTGCTTCAAGTCTTTAGAGGAATAAAAGGGAAGTGAGAGA 474

RESULT 13  
ABL98260/c  
ID ABL98260 standard; DNA; 3160 BP.  
XX  
AC ABL98260;  
XX  
DT 21-JUN-2002 (first entry)  
XX  
DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2912.  
XX  
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disease; infection; cytostatic; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155317-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US0001329.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.

```
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233297P.
PR 14-SEP-2000; 2000US-0233298P.
PR 14-SEP-2000; 2000US-0233299P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244674P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246612P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful
PT for preventing, diagnosing and/or treating testicular cancer.
XX
XX Disclosure; SEQ ID NO 2912; 766pp; English.
XX
XX The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a DNA encoding a
CC protein fragment of the invention
XX
XX Sequence 3160 BP; 861 A; 761 C; 855 G; 683 T; 0 U; 0 Other;
Query Match 29.4%; Score 684.2; DB 4; Length 3160;
Best Local Similarity 97.9%; Pred. No. 1.5e-166;
Matches 704; Conservative 0; Mismatches 13; Indels 2; Gaps 1;
QY 1602 CCTTCTGGTGGAATCTGGGCTCTGCTATTCTCAGCTGAGGATAGCTTCTCAGACCAAGGACCTGGAT 1661
Db 1190 CCTTAAAGGTGAATCTGGGCTCTGCTATTCTCAGCTGAGGATAGCTTCTCAGACCAAGGACCTGGAT 1131
QY 1662 ACCTCTTCTATTACGGTCCCGGCTCCAGGAGGATAGCTCCCAATGGGATGGGCCAC 1721
Db 1130 ACCTCTTCTATTACGGTCCCGGCTCCAGGAGGATAGCTCCCAATGGGATGGGCCAC 1071
QY 1722 TGAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1781
Db 1070 TGAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1011
QY 1782 GACAGGCAATCAAGGCTGAGCAACCAAAAGAGTGCCCAATATGGCTTTTCTACCTGTA 1841
```

Db 1010 GACAGCAATCAAGCCTGAGCAACCAAGGAGTGCCCATATGGCTTTCTACCTGTA 951  
 QY 1842 ACATGCACATAGAGCCATGCGCTAGATTATATAATACCAAGAGAGTTCTATTTTGA 1901  
 Db 950 ACATGCACATAGAGCCATGCGCTAGATTATATAATACCAAGAGAGTTCTATTTTGA 891  
 QY 1902 AAGACTGCAAAAAGGAGGAAAAAACCCTTCAAAAAGCGCCCTTAAGTCAACGCTCCATT 1961  
 Db 890 AAGACTGCAAAAAGGAGGAAAAAACCCTTCAAAAAGCGCCCTTAAGTCAACGCTCCATT 831  
 QY 1962 GACTGAAGACAGTCCCTATCTATAGAGGGTTGAGCTTTCTTCTCTTGGGTTGAGGAG 2021  
 Db 830 GACTGAAGACAGTCCCTATCTATAGAGGGTTGAGCTTTCTTCTCTTGGGTTGAGGAG 771  
 QY 2022 ACCAGGTCCTCTTATCTCTTCTAGCGGTCTGCGCTCTGCTGTTGAGGATCG 2081  
 Db 770 ACCAGGTCCTCTTATCTCTTCTAGCGGTCTGCGCTCTGCTGTTGAGGATCG 711  
 QY 2082 GCAACAGGCTACCCCTGAGGTCCTCATGTGCCATGAGTGTGCACAAATGTGCT 2141  
 Db 710 GCAACAGGCTACCCCTGAGGTCCTCATGTGCCATGAGTGTGCACA--CATGCATGTGCT 653  
 QY 2142 GTGTATGTGTAATGTGAGAAAAACACAGCCCTCTTTTCAAGAGAAAGGGGCTGAGGT 2201  
 Db 652 GTGTATGTGTAATGTGAGAGACACAGCCCTCTTTTCAAGAGAAAGGGGCTGAGGT 593  
 QY 2202 GCCAGCTGTCTGCTGTTAGGGTTAGGGGTCGGGCTCCCTTCCAGGGCCAGGAAGCAGG 2261  
 Db 592 GCCAGCTGTCTGCTGTTAGGGTTAGGGGTCGGGCTCCCTTCCAGGGCCAGGAAGCAGG 533  
 QY 2262 TTCCCTCTCTGCTGCTGCTTGTCAAGTCTTAGAGGAAATAAAAAGGGAAGTGAGAAA 2320  
 Db 532 TTCCCTCTCTGCTGCTGCTTGTCAAGTCTTAGAGGAAATAAAAAGGGAAGTGAGAGA 474

RESULT 14  
 AAL05397/c  
 ID AAL05397 standard; DNA; 7709 BP.  
 XX  
 AC AAL05397;  
 XX  
 DT 21-NOV-2001 (first entry)  
 XX  
 DE Human reproductive system related antigen DNA SEQ ID NO: 8085.  
 XX  
 KW Human; reproductive system related antigen; reproductive system disorder;  
 KW cancer; gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155320-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US0001339.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465570/50.  
XX  
XX Isolated nucleic acid molecule encoding a reproductive system antigen is  
XX used in preventing, treating or ameliorating a medical condition.  
XX  
XX Disclosure; SEQ ID NO 8085; 1297pp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
XX number of human reproductive system related antigens. These can be used  
XX in the prevention and treatment of reproductive system disorders,  
XX including cancer. The present sequence is a genomic sequence encoding a  
XX protein of the invention  
XX  
XX SQ Sequence 7709 BP; 2035 A; 1878 C; 2157 G; 1639 T; 0 U; 0 Other;  
Query Match 29.4%; Score 684.2; DB 4; Length 7709;  
Best Local Similarity 97.9%; Pred. No. 2.2e-166;  
Matches 704; Conservative 13; Mismatches 13; Indels 2; Gaps 1;  
PR 1602 CCTCTGGTGAATCGGCTCCCTGCTATTCTCACTCTGGGATTCCTGTGCTTCCT 1661  
DB 1188 CCCTAAAGTGAATCTGGGCTCTGCTATTCTCACTCTGGGATTCCTGTGCTTCCT 1129

QY 1662 ACCTCTTCTATTACCGTCCCGGCTCCAGCAGGAGTAGCGCGCAATGGGATGGGCCAC 1721  
DB 1128 ACCTCTTCTATTACCGTCCCGGCTCCAGCAGGAGTAGCGCGCAATGGGATGGGCCAC 1069  
QY 1722 TGAAGGTGCTTACCGGCTCTGAGGTGACCGCATAGACTTTCAGACCAAGGAGCCTGGAT 1781  
DB 1068 TGAAGGTGCTTACCGGCTCTGAGGTGACCGCATAGACTTTCAGACCAAGGAGCCTGGAT 1009  
QY 1782 GACAGGCAATCAAGGCTGAGCAACCAAAAGGAGTGCCCATATGCTTTTCTACCTGTA 1841  
DB 1008 GACAGGCAATCAAGGCTGAGCAACCAAAAGGAGTGCCCATATGCTTTTCTACCTGTA 949  
QY 1842 ACATGCACATAGAGCCATGGCGTAGATTTTATAATACCAAGAGAAGTTCTATTTTCTGA 1901  
DB 948 ACATGCACATAGAGCCATGGCGTAGATTTTATAATACCAAGAGAAGTTCTATTTTCTGA 889  
QY 1902 AAGACTGCAAAAAGGAGGAAAAAAACCTTCAAAAACGCCCCCTAAGTCAACGCTCCATT 1961  
DB 888 AAGACTGCAAAAAGGAGGAAAAAAACCTTCAAAAACGCCCCCTAAGTCAACGCTCCATT 829  
QY 1962 GACTGAAGACAGTCCCTATCCTAGAGGGTTGAGCTTCTTCTCTCTGGGTTGGAGGAG 2021  
DB 828 GACTGAAGACAGTCCCTATCCTAGAGGGTTGAGCTTCTTCTCTCTGGGTTGGAGGAG 769  
QY 2022 ACCAGGGTGCTCTTATCTCTCTAGCGGTCTGCTCTGCTGCTCTCTCTCTCTCTCTCTCT 2081  
DB 768 ACCAGGGTGCTCTTATCTCTCTAGCGGTCTGCTCTGCTGCTCTCTCTCTCTCTCTCTCT 709  
QY 2082 GCAAAACAGGCTACCCCTGAGGTCCCATGTGTCATGAGTGTGCACAAATGTAATGTGTCT 2141  
DB 708 GCAAAACAGGCTACCCCTGAGGTCCCATGTGTCATGAGTGTGCACAAATGTAATGTGTCT 651  
QY 2142 GTGTATGTGAATGTGAGAAAAACACAGCCCTCCTTTCAGAAAGGAGGCGCTGAGGT 2201  
DB 650 GTGTATGTGAATGTGAGAGAGACACAGCCCTCCTTTCAGAAAGGAGGCGCTGAGGT 591  
QY 2202 GCCAGCTGTGCTCTGGGTTAGGGGTGCGGCGGCTTCCAGGGCCAGGAAGGCGAGG 2261  
DB 590 GCCAGCTGTGCTCTGGGTTAGGGGTGCGGCGGCTTCCAGGGCCAGGAAGGCGAGG 531  
QY 2262 TTCCCTCTCTGGTCTGCTGCTTCAAGTCTTACAGGAAATAAAAAGGAAAGTGAAGAA 2320  
DB 530 TTCCCTCTCTGGTCTGCTGCTTCAAGTCTTACAGGAAATAAAAAGGAAAGTGAAGAA 472  
RESULT 15  
ABL98259/c  
ID ABL98259 standard; DNA; 7709 BP.  
XX AC ABL98259;  
XX XX  
XX 21-JUN-2002 (first entry)  
XX DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2911.  
XX XX  
XX KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
XX KW reproductive system disorder; urinary system disorder; gene therapy;  
XX KW cardiovascular disorder; respiratory disorder; neurological disorder;  
XX XX gastrointestinal disease; infection; cytostatic; gene; ds.  
XX OS Homo sapiens.  
XX XX  
XX PN WO200155317-A2.  
XX XX  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US001329.  
XX XX  
XX PR 31-JAN-2000; 2000US-0179065P.  
XX PR 04-FEB-2000; 2000US-0180628P.  
XX PR 24-FEB-2000; 2000US-0184664P.  
XX PR 02-MAR-2000; 2000US-0186350P.



PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225577P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234397P.  
PR 25-SEP-2000; 2000US-0234398P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251989P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483232/52.

Nucleic acids encoding 973 human testicular antigen polypeptides, useful for preventing, diagnosing and/or treating testicular cancer.

Disclosure; SEQ ID NO 2911; 766pp; English.

The present invention provides the protein and coding sequences of 973 human testicular antigens, and fragments of their genomic sequences. The

CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention

XX  
SQ Sequence 7709 BP; 2035 A; 1878 C; 2157 G; 1639 T; 0 U; 0 Other;

	Query Match	29.4%;	Score 684.2;	DB 4;	Length 7709;
	Best Local Similarity	97.9%;	Pred. No. 2.2e-166;		
	Matches 704;	Conservative 0;	Mismatches 13;	Indels 2;	Gaps 1;
Qy	1602	CCTCTCTGGGTGAATCTGGGCTCTCGCTATTCTCACTCTGGGATTCCTGTTCCTTCCT	1661		
Db	1188	CCCTAAGGTGAATCTGGGCTCTCGCTATTCTCACTCTGGGATTCCTGTTCCTTCCT	1129		
Qy	1662	ACCTCTTCTATTACCGTCCCGCTCCAGCAGAGTAGCCGCCAATGGGATGGGCCAC	1721		
Db	1128	ACCTCTTCTATTACCGTCCCGCTCCAGCAGAGTAGCCGCCAATGGGATGGGCCAC	1069		
Qy	1722	TGAAGGTCTTAGCGGCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT	1781		
Db	1068	TGAAGGTCTTAGCGGCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT	1009		
Qy	1782	GACAGGCAATCAAGGCTGAGCAACCAAAAGGAGTGCCCATATGGCTTTTCTACCTGTA	1841		
Db	1008	GACAGGCAATCAAGGCTGAGCAACCAAAAGGAGTGCCCATATGGCTTTTCTACCTGTA	949		
Qy	1842	ACATGCACATAGAGCATGGCGTAGATTATATAATACCAAGAGAGTTCTATTTTGTGA	1901		
Db	948	ACATGCACATAGAGCATGGCGTAGATTATATAATACCAAGAGAGTTCTATTTTGTGA	889		
Qy	1902	AGACTGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1961		
Db	888	AGACTGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	829		
Qy	1962	GACTGAAGACAGTCCCTATCTAGAGGGGTGAGCTTTCTCTCTCTCTCTCTCTCTCTCT	2021		
Db	828	GACTGAAGACAGTCCCTATCTAGAGGGGTGAGCTTTCTCTCTCTCTCTCTCTCTCTCT	769		
Qy	2022	ACCAGGTGCTCTTATCT	2081		
Db	768	ACCAGGTGCTCTTATCT	709		
Qy	2082	GCAACACAGCTACCCCTGAGGTCCCATGTGCGATGCTGCACACATGCAATGTGTCT	2141		
Db	708	GCAACACAGCTACCCCTGAGGTCCCATGTGCGATGCTGCACACATGCAATGTGTCT	651		
Qy	2142	GTGTATGTGTAATGTGAG	2201		
Db	650	GTGTATGTGTAATGTGAG	591		
Qy	2202	GCCAGCTGTCTCTGGGTAGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGGGG	2261		
Db	590	GCCAGCTGTCTCTGGGTAGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGGGG	531		
Qy	2262	TTCCCTCTCTGGGTCTGCTTGCAGCTTTAGAGGAGAGAGAGAGAGAGAGAGAGAG	2320		
Db	530	TTCCCTCTCTGGGTCTGCTTGCAGCTTTAGAGGAGAGAGAGAGAGAGAGAGAGAG	472		

Search completed: April 10, 2004, 20:05:45

Job time : 984.576 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:11:52 ; Search time 80.7618 Seconds  
(without alignments)  
10733.545 Million cell updates/sec

Title: US-09-743-825-10  
Perfect score: 20  
Sequence: 1 gaccgcataagactcttcaga 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl : \*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sta.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_ot.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sta.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	20	100.0	321	11	G22380	human STS W
C 2	20	100.0	457	6	AX407361	Sequence
C 3	20	100.0	2093	6	AX880371	AX880371 Sequence
C 4	20	100.0	2093	6	BD158342	BD158342 Primer fo
C 5	20	100.0	2093	9	AK027732	AK027732 Homo sapi
C 6	20	100.0	2233	6	AX768966	AX768966 Sequence
C 7	20	100.0	2326	9	AF045584	AF045584 Homo sapi
C 8	20	100.0	2373	6	BD063631	BD063631 Human pro
C 9	20	100.0	2375	9	BC001639	BC001639 Homo sapi
C 10	20	100.0	2525	9	AB103033	AB103033 Homo sapi
C 11	20	100.0	136678	2	AP001332	AP001332 Homo sapi
C 12	20	100.0	143899	2	AC013606	AC013606 Homo sapi
C 13	20	100.0	143899	6	AX411543	AX411543 Sequence
C 14	20	100.0	165434	2	AC087505	AC087505 Homo sapi
C 15	20	100.0	178000	9	AP002893	AP002893 Homo sapi
C 16	20	100.0	187566	2	AC015685	AC015685 Homo sapi
C 17	18.4	92.0	248825	2	AC132054	AC132054 Rattus no
C 18	18.4	92.0	260186	2	AC111920	AC111920 Rattus no
C 19	17.4	87.0	181595	2	AC136121	AC136121 Rattus no
C 20	17.4	87.0	222814	10	AC134591	AC134591 Mus muscu
C 21	16.8	84.0	2337	9	AB070000	AB070000 Macaca fa
C 22	16.8	84.0	38826	9	AC005356	AC005356 Homo sapi
C 23	16.8	84.0	59866	2	AC100572	AC100572 Mus muscu
C 24	16.8	84.0	110000	2	AC095743	Continuation (2 of
C 25	16.8	84.0	156075	10	AL844157	AL844157 Mouse DNA
C 26	16.8	84.0	159818	2	AC141050	AC141050 Rattus no
C 27	16.8	84.0	163511	9	AL442203	AL442203 Human DNA
C 28	16.8	84.0	168194	2	AC141475	AC141475 Mus muscu
C 29	16.8	84.0	172071	9	AC012676	AC012676 Homo sapi
C 30	16.8	84.0	175801	10	AC132464	AC132464 Mus muscu
C 31	16.8	84.0	180745	9	AL359183	AL359183 Human DNA
C 32	16.8	84.0	182679	9	AL365207	AL365207 Human DNA
C 33	16.8	84.0	183457	2	AC128106	AC128106 Rattus no
C 34	16.8	84.0	195582	2	AC105486	AC105486 Rattus no
C 35	16.8	84.0	196545	9	AC113390	AC113390 Homo sapi
C 36	16.8	84.0	212712	2	AC135641	AC135641 Mus muscu
C 37	16.8	84.0	213331	10	AC098745	AC098745 Mus muscu
C 38	16.8	84.0	214779	2	AC112769	AC112769 Rattus no
C 39	16.8	84.0	232067	2	AC121286	AC121286 Mus muscu
C 40	16.8	84.0	250620	2	AC137946	AC137946 Mus muscu
C 41	16.8	84.0	252675	2	AC116184	AC116184 Rattus no
C 42	16.8	84.0	253140	2	AC094055	AC094055 Rattus no
C 43	16.8	84.0	254575	2	AC127999	AC127999 Rattus no
C 44	16.8	84.0	255436	2	AC125922	AC125922 Rattus no
C 45	16.8	84.0	265985	2	AC136674	AC136674 Rattus no

ALIGNMENTS

RESULT 1  
G22380/c  
LOCUS human STS WI-17004, sequence tagged site.  
DEFINITION G22380  
ACCESSION G22380.1 GI:1342706  
VERSION STS; STS sequence; primer; sequence tagged site.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 321)  
AUTHORS Hudson,T.  
TITLE Whitehead Institute/MIT Center for Genome Research; Physically  
Mapped STSs

JOURNAL  
COMMENT

Unpublished (1995)

Contact: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
Whitehead Institute for Biomedical Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu

Primer A: TTTATAAATCTACGGCCATGGC  
Primer B: CTGGATGACAGGCAATCAAG  
STS size: 101  
PCR Profile:

## Presoak:

Denaturation:  
Annealing: 56 degrees C

Polymerization:

PCR Cycles: 35

Thermal Cycler:

Protocol:

Template: 10 ng

Primer: each 5 pm

dNTPs: each 4 nM

Taq Polymerase: 0.025 units/ul

Total Vol: 20 ul

## Buffer:

MgCl2: 1.5 mM

KCl: 50 mM

Tris-HCl: 10 mM

pH: 9.3

Derived from dbEST (genbank accession R00504).

FEATURES  
source

Location/Qualifiers

1..321

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/map="291.1 cR from top of ChrII linkage group"

35..135

35..56

complement(116..135)

## STS

primer\_bind

35..135

35..56

complement(116..135)

## ORIGIN

Query Match 100.0%; Score 20; DB 11; Length 321;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||

Db 164 GACCGCATAGACTTCTCAGA 145  
|||||

## RESULT 2

AX407361/c

LOCUS

DEFINITION

SEQUENCE 8 from Patent WO0229103.

ACCESSION

AX407361

VERSION

AX407361.1

GI:21440066

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.

Gene expression profiles in liver cancer

Patent: WO 0229103-A 8 11-APR-2002;

JOURNAL

GENE LOGIC INC (US)

FEATURES

source

Location/Qualifiers

1..457

/organism="Homo sapiens"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||

Db 144 GACCGCATAGACTTCTCAGA 125  
|||||

## RESULT 3

AX880371

LOCUS

DEFINITION

SEQUENCE 15276 from Patent EP1074617.

ACCESSION

AX880371

VERSION

AX880371.1

GI:40035107

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,

Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.

Primers for synthesizing full-length cDNA and their use

Patent: EP 1074617-A 15276 07-FEB-2001;

Research Association for Biotechnology (JP)

Location/Qualifiers

1..2093

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

77..1534

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAE90921.1"

/db\_xref="GI:40035108"

/translation="MLNLGFTIGSVLSATTLPLGILMDRPGPRVRLVGSACFTASC

TLMALASRDVEALSPLIFLALSLNPGFGICLTFTSLTPNNMGNRLSTLMALMTGSA

SSAITFPGIKLIYDAGVAVFVIMFTWSGLACLIPLNCTLNWPIEAFAPPEVNTTKI

KLGLALDHKVTGDLFYTHVTMGORLSOKAPSLDGSDAFMSQDVRGTSENLPERS

VPLKSLCSPTEFLWSLTMGTQRLIIFYAAVNNKLEYLVTCGQEHETNSQQQKVAE

TVGYSSVFGAMQLLCLTLNLLVFGITCLINNLHLQFVTFVLHTIVRGFFHSCGSL

RYCKIQKLTNAISAFITLNLGLSLISAVFALIQQLFMAMVGLKGEFPFVNLGLLLFSL

YAAVFPNHFETLTGLQSLISAVFALIQQLFMAMVGLKGEFPFVNLGLLLFSL

LLPSYLFYRRLQEQEYAAANGMPLKVLGSEVTA"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2093;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||

Db 1525 GACCGCATAGACTTCTCAGA 1544  
|||||

## RESULT 4

BD158342

LOCUS

DEFINITION

PRIMER for synthesizing full-length cDNA and use thereof.

ACCESSION

BD158342

VERSION

BD158342.1

GI:27864100

KEYWORDS

JP 2002191363-A/13185.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

(bases 1 to 2093)

AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayaashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
TITLE Primer for synthesizing full-length cDNA and use thereof  
JOURNAL Patent: JP 2002191363-A 13185 09-JUL-2002,  
HELIIX RESEARCH INSTITUTE  
COMMENT OS Homo sapiens (human)  
FN JP 2002191363-A/13185  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS (77) . . (1531) .  
Location/Qualifiers  
FEATURES  
source  
1. .2093  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 2093;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1525 GACCGCATAGACTTCTCAGA 1544  
RESULT 5  
AK027732 2093 bp mRNA linear PRI 01-AUG-2002  
LOCUS Homo sapiens cDNA FLJ14826 fis, clone OVARC1000850, highly similar  
DEFINITION To Homo sapiens PB39 mRNA.  
ACCESSION AK027732.1 GI:14042628  
VERSION oligo capping; fis (full insert sequence).  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Isogai,T., Ota,T., Hayaashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,  
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,  
Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,  
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,  
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,  
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,  
Nakamura,Y., Nagahara,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2093)  
Isogai,T. and Otsuki,T.  
AUTHORS Direct Submission  
TITLE Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
JOURNAL (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology; cDNA library construction,  
5'- & 3'-end one pass sequencing and clone selection; Helix  
Research Institute (supported by Japan Key Technology Center etc.)  
and Department of Virology, Institute of Medical Science,  
University of Tokyo.  
LOCATION/Qualifiers  
1. .2093

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="OVARC1000850"  
/tissue\_type="ovary, tumor tissue"  
/clone\_lib="OVARC1"  
/note="cloning vector: pME18SFL3"  
ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 2093;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1525 GACCGCATAGACTTCTCAGA 1544  
RESULT 6  
AX768966 2233 bp DNA linear PAT 02-JUL-2003  
LOCUS Sequence 83 from Patent WO02098917.  
DEFINITION AX768966  
ACCESSION AX768966  
VERSION AX768966.1 GI:32437147  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Guo,X., Fernandes,E., Li,L., Kekuda,R., Liu,Y., Leite,M.,  
Spytek,K.A., Ji,W., Casman,S.J., Boldog,F.L., Patturajan,M.,  
Vernet,C.A., Ballinger,R.A., Malvankar,U.M., Tchernev,V.T.,  
Blalock,A.D., Gusev,V.Y., Rastelli,L., Mezes,P.D., Ellerman,K.,  
Heyes,M., Herrmann,J.L., Shimkets,R.A., Ioine,N., Pena,C.E.,  
Shenoy,S.G., Taupier,R.J., Gerlach,V. and Gorman,L.  
TITLE Human proteins and nucleic acids encoding same  
JOURNAL Patent: WO 02098917-A 83 12-DEC-2002;  
Curagen Corporation (US)  
FEATURES  
source  
1. .2233  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 2233;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1654 GACCGCATAGACTTCTCAGA 1673  
RESULT 7  
AF045584 2326 bp mRNA linear PRI 27-AUG-1998  
LOCUS Homo sapiens PB39 mRNA, complete cds.  
DEFINITION AF045584  
ACCESSION AF045584  
VERSION AF045584.1 GI:3462514  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2326)  
AUTHORS Cole,K.A., Chuqui,R.F., Katz,K., Pack,S., Zhuang,Z., Cole,C.E.,  
Lyne,J.C., Linehan,W.M., Liotta,L.A. and Emmert-Buck,M.R.  
TITLE cDNA sequencing and analysis of POV1 (PB39): a novel gene  
up-regulated in prostate cancer  
JOURNAL Genomics 51 (2), 282-287 (1998)  
MEDLINE 98390192

```

9722952
PUBMED
REFERENCE
AUTHORS
  Cole, K.A., Chuaqui, R.F., Katz, K., Pack, S., Zhuang, Z., Cole, C.E.,
  Lyne, J.C., Linehan, M., Liotta, L.A. and Emmert-Buck, M.R.
TITLE
  Direct Submission
JOURNAL
  Submitted (31-JAN-1998) Laboratory of Pathology, NCI/NIH, 9000
  Rockville Pike, Bethesda, MD 20892, USA
FEATURES
  Location/Qualifiers
  source
    1..2326
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /chromosome="11"
      /map="11p11.1-p11.2"
      77..1756
        /note="up-regulated in prostate cancer"
        /codon_start=1
        /product="PB39"
        /protein_id="AAC33004.1"
        /db_xref="GI:3462515"
        /translation="MAPTQQAYRRRWMACTAVLENLFFSAVLLGWGSLIILKNEG
        FYSTCPAESSTNTQDQRWPQCQDEMLNGFTIGSFVLSATLPLGLMDRFG
        PRPVLVGSACFTACTMALASRDVEALSPLIALSLNGFGICLFTSLTPNMF
        GNRLSTLMALIGSYASSAIFPGIKLIYDAGVAVFMFTWGLACLIFFNCTLNWP
        IEAPFAPEEVNTKKIKLSGLALDHKVTGDLFYTHVTMTGQRLSKAPSLDGGDAFM
        SPQDVRGTSENLPERSVPLRSLCSPFLMSLLTGMGTQLRIIFVMAVNMKLEVLVT
        GQOHEHQOQKVAETRYGKQKLNALSAFTLNLGVGFGICLNLHLLQKVFV
        VLGDARDGATKSLRPRYKQKLNALSAFTLNLGVGFGICLNLHLLQKVFV
        LHTVYFPHGASGLSAAVFPNSHFGTLTGLSLISAVFALLOQLPFLPMVMPGLKGE
        PFWNLGGLLPSLLGFLPLSYLFYRRARLQOEYAANGMGLKVLGSGSEVTA"
  CDS
    100.0%; Score 20; DB 9; Length 2326;
    Best Local Similarity 100.0%; Pred. No. 2.5;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY
    1 GACCGCATAGACTTCTCAGA 20
    |||||||
  Db
    1747 GACCGCATAGACTTCTCAGA 1766
    |||||||
  RESULT 8
  BD063631
  LOCUS
    Human proteins having transmembrane domains and DNAs encoding these
    proteins.
  DEFINITION
    BD063631
  VERSION
    BD063631.1 GI:22609234
  KEYWORDS
    JP 2001508407-A/46.
  SOURCE
    Homo sapiens (human)
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 2373)
    Kato, S., Sekine, S., Kimura, T. and Kobayashi, M.
    Human proteins having transmembrane domains and DNAs encoding these
    proteins
  JOURNAL
    Patent: JP 2001508407-A 46 26-JUN-2001;
    SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
  COMMENT
    PS Homo sapiens (human)
    PN JP 2001508407-A/46
    PD 26-JUN-2001
    PF 07-NOV-1997 JP 1998522374
    PR 13-NOV-1996 JP 8/301429
    PI SISHI KATO, SHINGO SEKINE, TOMOKO KIMURA, MIDORI KOBAYASHI PC
    C12N15/12, C07K14/705, C12N15/10, C12N15/57, C12N9/48, C12N9/14, PC
    C12N15/55
    CC Strandedness: Double;
    CC Topology: Linear;
    FH Key Location/Qualifiers
    FT CDS 134..1813.
    FEATURES
    source
      1..2373

```

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
Query Match 100.0%; Score 20; DB 6; Length 2373;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY
  1 GACCGCATAGACTTCTCAGA 20
  |||||||
  Db
    1804 GACCGCATAGACTTCTCAGA 1823
    |||||||
  RESULT 9
  BC001639
  LOCUS
    Homo sapiens solute carrier family 43, member 1, mRNA (cDNA clone
    MGC:1944 IMAGE:2959372), complete cds.
  DEFINITION
    BC001639
  ACCESSION
    BC001639.2 GI:33876352
  VERSION
    MGC.
  KEYWORDS
    Homo sapiens (human)
  SOURCE
    Homo sapiens
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 2375)
    Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
    Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
    Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
    Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J.M., Hsieh, F.,
    Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
    Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
    Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,
    Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
    Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
    McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
    Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
    Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
    Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
    Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
    Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
    Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
    Butlerfield, Y.S., Krzywinski, M.I., Skalek, U., Smaluk, D.E.,
    Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
    Generation and initial analysis of more than 15,000 full-length
    human and mouse cDNA sequences
    Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  JOURNAL
    MEDLINE
    PUBMED
    12477932
  REFERENCE
    2 (bases 1 to 2375)
    Strausberg, R.
    Direct Submission
    Submitted (08-JAN-2001) National Institutes of Health, Mammalian
    Gene Collection (MGC), Cancer Genomics Office, National Cancer
    Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
    USA
  REMARK
    NIH-MGC Project URL: http://mgc.nci.nih.gov
    On Aug 19, 2003 this sequence version replaced gi:12804464.
    Contact: MGC help desk
    Email: gcgaps-r@mail.nih.gov
    Tissue Procurement: ATCC
    CDNA Library Preparation: Rubin Laboratory
    CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    DNA Sequencing by: Institute for Systems Biology
    http://www.systemsbio.org
    contact: amadan@systemsbio.org
    Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
    Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
    Clone distribution: MGC clone distribution information can be found
    through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
    Series: IRAL Plate: 2 Row: c Column: 12

```

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505970.

FEATURES  
source Location/Qualifiers

1..2375  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:1944 IMAGE:2959372"  
/issue\_type="Colon, adenocarcinoma"  
/clone\_lib="NIH MGC\_15"  
/lab\_host="DH10B-R\_1"  
/note="vector: pOTB7"  
1..2375  
/gene="SLC43A1"  
/note="synonyms: PB39, R00504, POV1"  
/db\_xref="LocusID:8501"  
/db\_xref="MIM:603733"  
119..1798  
/codon\_start=1  
/product="solute carrier family 43, member 1"  
/protein\_id="AAH01639.1"  
/db\_xref="GI:12804465"  
/db\_xref="LocusID:8501"  
/translation="MPTLQAYRRRWMACTAVLENLFFSAVLWGSLILKNEG  
FVSTCPAESSTNTTQDQRWPGCDQDDMLNLGFTIGSFVLSATTLPLGLMDRFG  
PRPRLVGSACFTASCTLMALASRDVEALSPILFLALSLNGFGGICLFTSLTLNMF  
GNRLSTLMALMIGSVASSAITPFGIKLIYDAGVAFVIMFTWSGLACLI FNCTLNWP  
IEAFPAPBEVNTYTKIKLSGLADHKVTDLPYHTVTMGQRLSKAPSLDGSDAFM  
SQDVRGTSENLPSVPLKSLCSPTFLWSLTTGMQTLRIIFMAAVNKMELVLT  
GQEHETNEQQKVAETVGFYSVFGAMQLLCTCLPLIGYIMDRIKDCVDAFTQGT  
VLGDARDGVATKIRPYCKIQKLNATSAFTLNLVGFGITCLINNLHQFTVFV  
LHTVIRGPFHSGSLYAAVFPNSHFGTLTGQSLISAVFALLOQLPMAMVGLKGE  
PFWNLGLLLPSLLGFLPSLYFYRRARLQQEYAANGMGLKVLGSSEVTA"

## gene

gene  
CDS

## CDS

ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 2375;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
DB 1789 GACCGCATAGACTTCTCAGA 1808

RESULT 10  
AB103033  
LOCUS Homo sapiens lat3 mRNA for L-type amino acid transporter 3, complete cds.  
DEFINITION AB103033  
ACCESSION AB103033.1 GI:38015937  
VERSION Homo sapiens (human)  
KEYWORDS Homo sapiens  
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ORGANISM  
REFERENCE  
AUTHORS Babu, E., Kanai, Y., Chairoungdua, A., Kim, D. K., Iribe, Y., Tangrongsup, S., Jutabha, P., Li, Y., Ahmed, N., Sakamoto, S., Anzai, N., Nagamori, S. and Endou, H.  
TITLE Identification of a novel system L amino acid transporter structurally distinct from heterodimeric amino acid transporters  
J. Biol. Chem. 278 (44), 43838-43845 (2003)

JOURNAL MEDLINE  
PUBMED 22940096  
REFERENCE 2 (bases 1 to 2525)  
AUTHORS Kanai, Y.  
TITLE Direct Submission  
JOURNAL Submitted (07-FEB-2003) Yoshikatsu Kanai, Kyorin University School of Medicine, Department of Pharmacology and Toxicology; 6-20-2 Shinkawa, Mitaka, Tokyo 181-8611, Japan  
(E-mail: ykanai@kyorin-u.ac.jp, Tel: 81-422-47-5511 (ex. 3453), Fax: 81-422-79-1321)

FEATURES  
source Location/Qualifiers

1..2525  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="FLC4"  
/cell\_type="hepatoma"  
/tissue\_type="liver"  
/clone\_lib="psPORT1 FLC4 human hepatoma cell xDNA"  
/country="Japan"  
1..2525  
/gene="lat3"  
257..1936  
/gene="lat3"  
/function="system L amino acid transporter"  
/codon\_start=1  
/product="L-type amino acid transporter 3"  
/protein\_id="BAD00152.1"  
/db\_xref="GI:38015938"  
/translation="MPTLQAYRRRWMACTAVLENLFFSAVLWGSLILKNEG  
FVSTCPAESSTNTTQDQRWPGCDQDDMLNLGFTIGSFVLSATTLPLGLMDRFG  
PRPRLVGSACFTASCTLMALASRDVEALSPILFLALSLNGFGGICLFTSLTLNMF  
GNRLSTLMALMIGSVASSAITPFGIKLIYDAGVAFVIMFTWSGLACLI FNCTLNWP  
IEAFPAPBEVNTYTKIKLSGLADHKVTDLPYHTVTMGQRLSKAPSLDGSDAFM  
SQDVRGTSENLPSVPLKSLCSPTFLWSLTTGMQTLRIIFMAAVNKMELVLT  
GQEHETNEQQKVAETVGFYSVFGAMQLLCTCLPLIGYIMDRIKDCVDAFTQGT  
VLGDARDGVATKIRPYCKIQKLNATSAFTLNLVGFGITCLINNLHQFTVFV  
LHTVIRGPFHSGSLYAAVFPNSHFGTLTGQSLISAVFALLOQLPMAMVGLKGE  
PFWNLGLLLPSLLGFLPSLYFYRRARLQQEYAANGMGLKVLGSSEVTA"

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 2525;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
DB 1927 GACCGCATAGACTTCTCAGA 1946

RESULT 11  
AP001332/c

LOCUS Homo sapiens chromosome 11 clone CMB9-114K6 map 11q12, WORKING  
DEFINITION AP001332.3 GI:9927272  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
Homo sapiens 136,678 genomic DNA of 11q12  
Published Only in Database (2000)

JOURNAL  
AUTHORS 2 (bases 1 to 136678)  
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-MAR-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555, Japan  
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923, Fax: 81-42-778-9924)  
On Aug 26, 2000 this sequence version replaced gi:8117255.

## COMMENT

----- Genome Center  
Center: RIKEN Genomic Sciences Center (GSC)  
Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information

Center project name: HumDraft11  
 Center clone name: CMB9-114K6  
 ----- Summary Statistics -----  
 Sequencing vector: PCR products; 100% of reads  
 Chemistry: Dye-terminator ET-amersham; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 122948 bases at least Q40  
 Consensus quality: 129078 bases at least Q30  
 Consensus quality: 13272 bases at least Q20  
 Insert size: 134078; sum-of-contigs  
 Quality coverage: 10.08x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 27 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 18878 contig of 18878 bp in length
18979 40676 contig of 21698 bp in length
40777 57862 contig of 17086 bp in length
57963 68848 contig of 10886 bp in length
68949 77918 contig of 8970 bp in length
78019 86066 contig of 8048 bp in length
86167 92722 contig of 6556 bp in length
92823 98264 contig of 5442 bp in length
98365 103665 contig of 5301 bp in length
103766 107189 contig of 3424 bp in length
107290 110555 contig of 3266 bp in length
110656 112554 contig of 1899 bp in length
112655 114861 contig of 2207 bp in length
114962 116468 contig of 1507 bp in length
116569 118769 contig of 2201 bp in length
118870 120846 contig of 1977 bp in length
120947 124644 contig of 1820 bp in length
124745 126582 contig of 1838 bp in length
126683 128325 contig of 1643 bp in length
128426 129436 contig of 1011 bp in length
129537 130773 contig of 1237 bp in length
130874 132008 contig of 1135 bp in length
132109 133187 contig of 1079 bp in length
133288 134426 contig of 1139 bp in length
134527 135528 contig of 1002 bp in length
135629 136678 contig of 1050 bp in length.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 18878 contig of 18878 bp in length
18979 40676 contig of 21698 bp in length
40777 57862 contig of 17086 bp in length
57963 68848 contig of 10886 bp in length
68949 77918 contig of 8970 bp in length
78019 86066 contig of 8048 bp in length
86167 92722 contig of 6556 bp in length
92823 98264 contig of 5442 bp in length
98365 103665 contig of 5301 bp in length
103766 107189 contig of 3424 bp in length

```

```

* 107190 107289: gap of 100 bp
* 107290 110555: contig of 3266 bp in length
* 110556 110655: gap of 100 bp
* 110656 112554: contig of 1899 bp in length
* 112555 112654: gap of 100 bp
* 112655 114861: contig of 2207 bp in length
* 114862 114961: gap of 100 bp
* 114962 116468: contig of 1507 bp in length
* 116469 116568: gap of 100 bp
* 116569 118769: contig of 2201 bp in length
* 118770 118869: gap of 100 bp
* 118870 120846: contig of 1977 bp in length
* 120847 120946: gap of 100 bp
* 120947 122766: contig of 1820 bp in length
* 122767 122866: gap of 100 bp
* 122867 124644: contig of 1778 bp in length
* 124645 124744: gap of 100 bp
* 124745 126582: contig of 1838 bp in length
* 126583 126682: gap of 100 bp
* 126683 128325: contig of 1643 bp in length
* 128326 128425: gap of 100 bp
* 128426 129436: contig of 1011 bp in length
* 129437 129536: gap of 100 bp
* 129537 130773: contig of 1237 bp in length
* 130774 130873: gap of 100 bp
* 130874 132008: contig of 1135 bp in length
* 132009 132108: gap of 100 bp
* 132109 133187: contig of 1079 bp in length
* 133188 133287: gap of 100 bp
* 133288 134426: contig of 1139 bp in length
* 134427 134526: gap of 100 bp
* 134527 135528: contig of 1002 bp in length
* 135529 135628: gap of 100 bp
* 135629 136678: contig of 1050 bp in length.

FEATURES
      source
      1..136678
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="11"
        /map="11q12"
        /clone="CMB9-114K6"
      1..18878
        /note="assembly_fragment"
      18979..40676
        /note="assembly_fragment"
      40777..57862
        /note="assembly_fragment"
      57963..68848
        /note="assembly_fragment"
      68949..77918
        /note="assembly_fragment"
      78019..86066
        /note="assembly_fragment"
      86167..92722
        /note="assembly_fragment"
      92823..98264
        /note="assembly_fragment"
      98365..103665
        /note="assembly_fragment"
      103766..107189
        /note="assembly_fragment"
      107290..110555
        /note="assembly_fragment clone_end:T7 vector_side:right"
      110556..112554
        /note="assembly_fragment"
      112555..114861
        /note="assembly_fragment"
      114962..116468
        /note="assembly_fragment"
      116569..118769
        /note="assembly_fragment"
      118870..120846
        /note="assembly_fragment"

```



```

misc_feature      /note="assembly_fragment"
120947..122766
/note="assembly_fragment"
122867..124644
/note="assembly_fragment"
124745..126582
/note="assembly_fragment"
126683..128325
/note="assembly_fragment"
128426..129436
/note="assembly_fragment"
129537..130773
/note="assembly_fragment"
130874..132008
/note="assembly_fragment"
132109..133187
/note="assembly_fragment"
133288..134426
/note="assembly_fragment"
134527..135528
/note="assembly_fragment"
135629..136678
/note="assembly_fragment"

```

Query Match 100.0%; Score 20; DB 2; Length 136678;

Best Local Similarity 100.0%; Pred. No. 4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GACCGCATGACTTCTCAGA 20
    |||||
DB 70834 GACCGCATGACTTCTCAGA 70815

```

#### RESULT 12

```

AC013606/c
LOCUS      AC013606      143899 bp      DNA      linear      HTG 12-MAR-2000
DEFINITION Homo sapiens clone RP11-10E21, WORKING DRAFT SEQUENCE, 14 unordered
pieces.
ACCESSION  AC013606
VERSION    AC013606.5 GI:7230045
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 143899)
AUTHORS    Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE      Homo sapiens, clone RP11-10E21
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 143899)
AUTHORS    Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
            Baldwin, J., Barna, N., Beckerly, K., Boguslavsky, L., Boukhgaiter, B.,
            Brown, A., Castle, A., Collangelo, M., Collins, S., Collymore, A.,
            Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
            Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
            Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
            Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
            Lechoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
            McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
            Mortenson, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
            Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
            Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
            Tesfaye, S., Turrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
            Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
            Direct Submission
            Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT    On Mar 12, 2000 this sequence version replaced gi:6850474.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research

```

```

Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3055
Center clone name: 10_E_21
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 13858 bases at least Q40
Consensus quality: 141075 bases at least Q30
Consensus quality: 141906 bases at least Q20
Insert size: 176000; agarose-fp
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 5.8 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2043: contig of 2043 bp in length
* 2044 2143: gap of 100 bp
* 2144 6624: contig of 4481 bp in length
* 6625 6724: gap of 100 bp
* 6725 14152: contig of 7428 bp in length
* 14153 14252: gap of 100 bp
* 14253 19511: contig of 5259 bp in length
* 19512 19611: gap of 100 bp
* 19612 22594: contig of 2983 bp in length
* 22595 22694: gap of 100 bp
* 22695 27824: contig of 5130 bp in length
* 27825 27924: gap of 100 bp
* 27925 34952: contig of 7028 bp in length
* 34953 35052: gap of 100 bp
* 35053 40782: contig of 5730 bp in length
* 40783 40882: gap of 100 bp
* 40883 48999: contig of 8117 bp in length
* 49000 49099: gap of 100 bp
* 49100 62883: contig of 13784 bp in length
* 62884 62983: gap of 100 bp
* 62984 75527: contig of 12544 bp in length
* 75528 75627: gap of 100 bp
* 75628 87943: contig of 12316 bp in length
* 87944 88043: gap of 100 bp
* 88044 111029: contig of 22386 bp in length
* 111030 111229: gap of 100 bp
* 111230 143899: contig of 32770 bp in length.
*
* Location/Qualifiers
* 1..143899
*   /organism="Homo sapiens"
*   /mol_type="genomic DNA"
*   /db_xref="taxon:9606"
*   /clone="RP11-10E21"
*   /clone_lib="RPC1-11 Human Male BAC"
* 1..2043
*   /note="assembly_fragment"
* 2144..6624
*   /note="assembly_fragment"
* 6725..14152
*   /note="assembly_fragment"
* 14253..19511
*   /note="assembly_fragment"
*   clone_end:77
*   vector_side:right
* 19612..22594
*   /note="assembly_fragment"
*   clone_end:SP6

```

```

misc_feature      vector side:left"
22695_-27824
/note="assembly_fragment"
misc_feature      27925_.34952
/note="assembly_fragment"
misc_feature      35053_.40782
/note="assembly_fragment"
misc_feature      40883_.48999
/note="assembly_fragment"
misc_feature      49100_.62883
/note="assembly_fragment"
misc_feature      62984_.75527
/note="assembly_fragment"
misc_feature      75628_.87943
/note="assembly_fragment"
misc_feature      8044_.111029
/note="assembly_fragment"
misc_feature      11130_.143899
/note="assembly_fragment"

ORIGIN
Query Match      100.0%; Score 20; DB 2; Length 143899;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20
Db 38586 GACCGCATAGACTTCTCAGA 38567

RESULT 13
AX411543/c AX411543 143899 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 15 from Patent W00229059.
DEFINITION AX411543
ACCESSION AX411543
VERSION AX411543.1 GI:21444137
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Sah,D.W.Y., Cate,R.L. and Strittmatter,S.M.
Nogo receptor homologs
Patent: WO 0229059-A 15 11-APR-2002;
BIOGEN INC (US)
FEATURES
source
1..143899
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 143899;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20
Db 38586 GACCGCATAGACTTCTCAGA 38567

RESULT 14
AC087505
LOCUS 165434 bp DNA linear HTG 07-JUL-2001
DEFINITION Homo sapiens chromosome 11 clone RP11-676P16 map 11, WORKING DRAFT
SEQUENCE, 12 unordered pieces.
ACCESSION AC087505
VERSION AC087505.4 GI:14626336
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

---

```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 165434)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-676P16
Unpublished
2 (bases 1 to 165434)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Petterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,
Sougnaz,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 7, 2001 this sequence version replaced gi:14029937.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12298
Center clone name: 676_P.16
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 160902 bases at least Q40
Consensus quality: 163021 bases at least Q30
Consensus quality: 163793 bases at least Q20
Insert size: 164000; agarose-fp
Insert size: 164334; sum-of-contigs
Quality coverage: 12.9 in Q20 bases; agarose-fp
Quality coverage: 12.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 15621: contig of 15621 bp in length
* 15622 15721: gap of 100 bp
* 15722 16822: contig of 1101 bp in length
* 16823 16922: gap of 100 bp
* 16923 18786: contig of 1864 bp in length
* 18787 18886: gap of 100 bp
* 18887 20784: contig of 1898 bp in length
* 20785 20884: gap of 100 bp
* 20885 22630: contig of 1746 bp in length
* 22631 22730: gap of 100 bp
* 22731 25643: contig of 2913 bp in length

```

\* 25644 25743: gap of 100 bp  
 \* 25744 27851: contig of 2108 bp in length  
 \* 27852 27952: gap of 100 bp  
 \* 27953 108833: contig of 80882 bp in length  
 \* 108834 108933: gap of 100 bp  
 \* 108934 122115: contig of 13182 bp in length  
 \* 122116 122215: gap of 100 bp  
 \* 122216 145049: contig of 22834 bp in length  
 \* 145050 145149: gap of 100 bp  
 \* 145150 162807: contig of 17858 bp in length  
 \* 162808 162907: gap of 100 bp  
 \* 162908 165434: contig of 2527 bp in length.

## FEATURES

source  
 1..165434  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11"  
 /clone="RP11-676P16"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 misc\_feature  
 1..15621  
 /note="assembly\_fragment  
 clone\_end:SP6  
 vector\_side:left"  
 misc\_feature  
 15722..16822  
 /note="assembly\_fragment"  
 misc\_feature  
 16923..18786  
 /note="assembly\_fragment"  
 misc\_feature  
 18887..20784  
 /note="assembly\_fragment"  
 misc\_feature  
 20885..22630  
 /note="assembly\_fragment"  
 misc\_feature  
 22731..25643  
 /note="assembly\_fragment"  
 misc\_feature  
 25744..27851  
 /note="assembly\_fragment"  
 misc\_feature  
 27952..108833  
 /note="assembly\_fragment"  
 misc\_feature  
 108934..122115  
 /note="assembly\_fragment"  
 misc\_feature  
 122216..145049  
 /note="assembly\_fragment"  
 misc\_feature  
 145150..162807  
 /note="assembly\_fragment"  
 misc\_feature  
 162908..165434  
 /note="assembly\_fragment  
 clone\_end:T7  
 vector\_side:right"

## ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 165434;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
 |||||

Db 144436 GACCGCATAGACTTCTCAGA 144455  
 |||||

## RESULT 15

AP002893/c 178000 bp DNA linear PRI 30-MAR-2002.  
 LOCUS  
 DEFINITION Homo sapiens genomic DNA, chromosome 11q clone:RP11-624G17,  
 complete sequence.  
 ACCESSION AP002893  
 VERSION AP002893.3 GI:19879819  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
 TITLE Homo sapiens genomic DNA  
 JOURNAL Published Only in Database (2000)  
 REFERENCE 2 (bases 1 to 178000)  
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-OCT-2000) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
 1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/  
 Tel:81-45-503-9111, Fax:81-45-503-9170)  
 COMMENT On Apr 1, 2002 this sequence version replaced gi:13469698.  
 FEATURES  
 Location/Qualifiers  
 1..178000  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11q"  
 /clone="RP11-624G17"

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 178000;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
 |||||

Db 55602 GACCGCATAGACTTCTCAGA 55583  
 |||||

Search completed: April 10, 2004, 17:33:05  
 Job time : 83.7618 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:11:52 ; Search time 88.838 Seconds  
(without alignments)  
10733.545 Million cell updates/sec

Title: US-09-743-825-7  
Perfect score: 22  
Sequence: 1 gcatgttacagtagaagcc 22

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_ot.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	321	11	G22380	G22380 human STS W
2	22	100.0	457	6	AX407361	AX407361 Sequence
3	22	100.0	2093	6	AX880371	AX880371 Sequence
4	22	100.0	2093	6	BD158342	BD158342 Primer fo
5	22	100.0	2093	9	AK027732	AK027732 Homo sapi
6	22	100.0	2233	6	AX768966	AX768966 Sequence
7	22	100.0	2326	9	AF045584	AF045584 Homo sapi
8	22	100.0	2373	6	BD063631	BD063631 Human pro
9	22	100.0	2375	9	BC001639	BC001639 Homo sapi
10	22	100.0	2525	9	AB103033	AB103033 Homo sapi
11	22	100.0	136678	2	AP001332	AP001332 Homo sapi
12	22	100.0	143899	2	AC013606	AC013606 Homo sapi
13	22	100.0	143899	6	AX411543	AX411543 Sequence
14	22	100.0	165434	2	AC087505	AC087505 Homo sapi
15	22	100.0	178000	9	AP002893	AP002893 Homo sapi
16	22	100.0	187566	2	AC015685	AC015685 Homo sapi
17	19.4	88.2	601	11	G88016	G88016 S208P659PF
18	19.4	88.2	192042	9	AC079955	AC079955 Homo sapi
19	19.4	88.2	243654	10	AL663082	AL663082 Mouse DNA
20	19	86.4	83709	2	AC144727	AC144727 Medicago
21	19	86.4	102421	2	AC146755	AC146755 Medicago
22	19	86.4	102858	2	AC144656	AC144656 Medicago
23	19	86.4	103039	8	AC139525	AC139525 Medicago
24	19	86.4	108694	2	AC135848	AC135848 Medicago
25	19	86.4	118705	8	AC126009	AC126009 Medicago
26	19	86.4	125883	8	AC124959	AC124959 Medicago
27	19	86.4	126687	2	AC131240	AC131240 Medicago
28	19	86.4	140418	2	AC137822	AC137822 Medicago
29	19	86.4	248040	2	AC137521	AC137521 Medicago
30	18.8	85.5	579	6	AX874028	AX874028 Sequence
31	18.8	85.5	579	6	BD154090	BD154090 Primer fo
32	18.8	85.5	171673	5	AL928790	AL928790 Zebrafish
33	18.8	85.5	181862	5	BX255897	BX255897 Zebrafish
34	18.4	83.6	47069	9	AC021101	AC021101 Homo sapi
35	18.4	83.6	156220	2	AC138738	AC138738 Mus muscu
36	18.4	83.6	192015	2	AC119223	AC119223 Mus muscu
37	18.4	83.6	199670	2	AC144802	AC144802 Mus muscu
38	18	81.8	230127	10	AL691481	AL691481 Mouse DNA
39	18	81.8	348498	1	BX571872	BX571872 Photorhab
40	18	81.8	349980	6	AX770910	AX770910 Sequence
41	17.8	80.9	2476	8	KLKLEU2G	X65545 K.lactis KL
42	17.8	80.9	43051	10	BX649238	BX649238 Mouse DNA
43	17.8	80.9	50374	2	AC015170	AC015170 Drosophila
44	17.8	80.9	63839	9	AL139394	AL139394 Human DNA
45	17.8	80.9	97065	2	AP000768	AP000768 Homo sapi

ALIGNMENTS

RESULT 1  
G22380 human STS WI-17004, sequence tagged site.  
LOCUS G22380 321 bp DNA linear STS 31-MAY-1996  
DEFINITION G22380  
ACCESSION G22380  
VERSION G22380.1 GI:1342706  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 321)  
AUTHORS Hudson, T.  
TITLE Whitehead Institute/MIT Center for Genome Research; Physically Mapped STSs

```

FEATURES
source
Location/Qualifiers
1. .457
/organism="Homo sapiens"

```

; Eutheria; Pri

; Eutheria; Pri

**AUTHORS**  
Ota, T., Isogai, T., Nishikawa, T., Hayaishi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
**TITLE**  
Primer for synthesizing full-length cDNA and use thereof  
**JOURNAL**  
Patent: JP 2002191363-A 13185 09-JUL-2002;  
HELIIX RESEARCH INSTITUTE  
**COMMENT**  
OS Homo sapiens (human)  
FN JP 2002191363-A/13185  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS (77) . (1531).  
Location/Qualifiers  
1. .2093  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

**ORIGIN**  
Query Match 100.0%; Score 22; DB 6; Length 2093;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
Db 1625 GCATGTTACAGGTAGAAAAGCC 1604

**FEATURES**  
source  
1. .2093  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

**ORIGIN**  
Query Match 100.0%; Score 22; DB 6; Length 2093;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
Db 1625 GCATGTTACAGGTAGAAAAGCC 1604

**RESULT 5**  
AK027732/c  
LOCUS  
DEFINITION Homo sapiens cDNA FLJ14826 fis, clone, OVARC1000850, highly similar to Homo sapiens PB39 mRNA.  
ACCESSION AK027732.1 GI:14042628  
VERSION AK027732.1  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Isogai, T., Ota, T., Hayaishi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2093)  
Isogai, T. and Otsuki, T.  
Direct Submission  
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - & 3' - end pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.  
Location/Qualifiers  
1. .2093

**FEATURES**  
source  
1. .2093  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="OVARC1000850"  
/tissue\_type="ovary, tumor tissue"  
/clone\_lib="OVARC1"  
/note="cloning vector: pME18SFL3"

**ORIGIN**  
Query Match 100.0%; Score 22; DB 9; Length 2093;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
Db 1625 GCATGTTACAGGTAGAAAAGCC 1604

**RESULT 6**  
AX768966/c  
LOCUS  
DEFINITION Sequence 83 from Patent WO02098917.  
ACCESSION AX768966  
VERSION AX768966.1 GI:32437147  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Guo, X., Fernandes, E., Li, L., Kekuda, R., Liu, Y., Leite, M., Spytek, K. A., Ji, W., Casman, S. J., Boldog, F. L., Patturajan, M., Vernet, C. A., Ballinger, R. A., Malyankar, U. M., Tchernev, V. T., Blalock, A. D., Gusev, V. Y., Rastelli, L., Mezes, P. D., Ellerman, K., Heyes, M., Herrmann, J. L., Shimkets, R. A., Ioine, N., Pena, C. E., Shenoy, S. G., Taupier, R. J., Gerlach, V. and Gorman, L.  
Human proteins and nucleic acids encoding same  
Patent: WO 02098917-A 83 12-DEC-2002;  
Curagen Corporation (US)  
Location/Qualifiers  
1. .2233  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

**ORIGIN**  
Query Match 100.0%; Score 22; DB 6; Length 2233;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
Db 1754 GCATGTTACAGGTAGAAAAGCC 1733

**RESULT 7**  
AF045584/c  
LOCUS  
DEFINITION Homo sapiens PB39 mRNA, complete cds.  
ACCESSION AF045584  
VERSION AF045584.1 GI:3462514  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2326)  
Cole, K. A., Chuqui, R. F., Katz, K., Pack, S., Zhuang, Z., Cole, C. E., Lyne, J. C., Linehan, W. M., Liotta, L. A. and Emmert-Buck, M. R.  
CDNA sequencing and analysis of POVI (PB39): a novel gene up-regulated in prostate cancer  
JOURNAL Genomics 51 (2), 282-287 (1998)  
MEDLINE 98390192

```

9722952
PUBMED 2 (bases 1 to 2326)
REFERENCE Cole, K.A., Chuahui, R.F., Katz, K., Pack, S., Zhuang, Z., Cole, C.E.,
AUTHORS Lyne, J.C., Linehan, M., Liotta, L.A. and Emmert-Buck, M.R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-1998) Laboratory of Pathology, NCI/NIH, 9000
Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1..2326
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p11.1-p11.2"
77..1756
/note="up-regulated in prostate cancer"
/codon_start=1
/product="PB39"
/protein_id="AAC33004.1"
/db_xref="GI:3462515"
/translation="NAPTQQAYRRRWMACTAVLENLFFPSAVLLGWGSLIILKNEG
FYSSCPASSNTTQDSQRRWPCQDQEMLNUGFTIGSFVLSATLPLGLMDRFG
PRPVLGASCTASCTLMALASRDVEALSPLIFLALSLNGFGGCTFTTSLTPNMF
GNLRTLMAIYKTSYASSAITPPGIKLIYDAGAFVIMFTWISGLACLIFFNCTLNWP
IEAPFAPEVNTYKKIKLSGLADHKVTGDLFYHTVTMGORLSOKAPSLDGSDAPM
SPQDVRGTSNLPERSVPLKSLCSPFLMSLLTGMGTQLRIIEYMAAVNKMLEVLT
GGQHEHTEQQQKVATVGFYSSVFGAMQLLCTCLPLIGYIMWRINKCDVATQGT
VIGDARDGATKTSIPRYCKIQKLNATSAFTLNLLLVGFGITCLNNKHQFVTFV
LHTIVRGFFHACGSLYAAVPSNHFGLTGLQSLISAVFALLOQPLPMVMVGLKGE
PFWNLGLLFLSLGFLPLPSLYFYRRARLQOEYAAANGMGLKVLGSGSEVTA"
ORIGIN
Query Match 100.0%; Score 22; DB 9; Length 2326;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCATGTTACAGGTAGAAAAGCC 22
Db 1847 GCATGTTACAGGTAGAAAAGCC 1826
RESULT 8
BD063631/c
LOCUS BD063631.1 GI:22609234
DEFINITION Human proteins having transmembrane domains and DNAs encoding these
proteins.
ACCESSION BD063631
VERSION BD063631.1
KEYWORDS JP 2001508407-A/46.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2373)
AUTHORS Kato, S., Sekine, S., Kimura, T. and Kobayashi, M.
TITLE Human proteins having transmembrane domains and DNAs encoding these
proteins
JOURNAL Patent: JP 2001508407-A 46 26-JUN-2001;
COMMENT SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
PS Homo sapiens (human)
ON JP 2001508407-A/46
PD 26-JUN-2001
PF 07-NOV-1997 JP 1998522374
PR 13-NOV-1996 JP 8/301429
PT SEISHI KATO, SHINGO SEKINE, TOMOKO KIMURA, MIDORI KOBAYASHI PC
C12N15/12, C07K14/705, C12N5/10, C12N15/57, C12N9/48, C12N9/14, PC
C12N15/55
CC Strandedness: Double;
CC Topology: Linear;
FT Key Location/Qualifiers
1..2373 Location/Qualifiers
1..2373
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```



This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505970.

FEATURES	Location/Qualifiers
source	1. .2375

gene	CDS
1. .2375	

```

CDS
119- .1798
/WW XLEI= MIN:003733

```

## ORIGIN

```
Query Match      100.0%; Score 22; DB 9; Length 2525;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 22: Conservative 0; Mismatches 0; Indels
```

QY 1 GCATGTTACAGGTAGAAAGCC 22  
227 GCATGTTACAGGTAGAAAGCC 2006

Accession	Sequence	Length	Result
U7	GCATGTTACAGGTAGAAAAGCC	1868	RESULT 11
Db	GCATGTTACAGGTAGAAAAGCC	1868	AP001332

LOCUS	AP001332	136678 bp	DNA linear	HTG 25-AUG-2000
DEFINITION	Homo sapiens chromosome 11 clone CMB9-114K6 map 11q12, WORKING DRAFT SEQUENCE. 27 unordered pieces.			

AB103033/C	AB103033	2525 bp	linear	PRI 29-OCT-2003	ACCESSION
	LOCUS				VERSION
	DEFINITION	Homo sapiens lat3 mRNA for L-type amino acid transporter 3, complete cds.			KEYWORDS
	ACCESSION	AB103033			SOURCE
					ORGANISM

AB103033.1 GI:38015937  
KEYWORDS

REFERENCE	1	Babu, E., Kanai, Y., Chairoungdua, A., Kim, D.K., Iribe, Y.	JOURNAL OF LIPID RESEARCH
AUTHORS			REFERENCE

TITLE	TITLE
Identification of a novel system L amino acid transporter	Identification of a novel system L amino acid transporter
structurally distinct from heterodimeric amino acid transporters	structurally distinct from heterodimeric amino acid transporters
ANZAI, N., NAGAMORI, S. and ENDOU, H.	ANZAI, N., NAGAMORI, S. and ENDOU, H.

JOURNAL  
Submitted (02-MAR-2000) Masahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan  
(E-mail: hattori@gsc.riken.go.jp, URL: <http://ngp.gsc.riken.go.jp/>,  
Tel: 81-42-778-9923, Fax: 81-42-778-9924)  
On Aug 26, 2000 this sequence version replaced at: 8117255.  
COMMENT

**TITLE**  
**JOURNAL**  
**Direct Submission**  
**Submitted (07-FEB-2003)** Yoshikatsu Kanai, Kyorin University School  
of Medicine, Department of Pharmacology and Toxicology, 5-20-2

Center code: RIKEN  
Web site: <http://hgp.gsc.riken.go.jp/>  
Contact: [hattori@gsc.riken.go.jp](mailto:hattori@gsc.riken.go.jp)  
-----Project Information-----

COMMENT

----- Genome Center  
Center: RIKEN Genomic Sciences Center (GSC)  
Center code: RIKEN  
Web site: <http://hgp.gsc.riken.go.jp/>  
Contact: [hattori@gsc.riken.go.jp](mailto:hattori@gsc.riken.go.jp)  
----- Protect Information -----

```

Center project name: HumDraft11
Center clone name: CMB9-114K6
----- Summary Statistics -----
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 122948 bases at least Q40
Consensus quality: 129078 bases at least Q30
Consensus quality: 132272 bases at least Q20
Insert size: 134078; sum-of-contigs
Quality coverage: 10.08x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
27 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.

1
18878 contig of 18878 bp in length
18979 40676 contig of 21698 bp in length
40777 57862 contig of 17086 bp in length
57963 68848 contig of 10886 bp in length
68949 77918 contig of 8970 bp in length
78019 86066 contig of 8048 bp in length
86167 92722 contig of 6556 bp in length
92823 98264 contig of 5442 bp in length
98365 103665 contig of 5301 bp in length
103666 107290 contig of 3266 bp in length
107290 110555 contig of 100 bp
110555 112554 contig of 100 bp
112554 114861 contig of 100 bp
114861 116468 contig of 1507 bp in length
116468 118870 contig of 100 bp
118870 120846 contig of 100 bp
120846 122948 contig of 100 bp
122948 124948 contig of 100 bp
124948 126948 contig of 100 bp
126948 128948 contig of 100 bp
128948 130948 contig of 100 bp
130948 132948 contig of 100 bp
132948 134948 contig of 100 bp
134948 136948 contig of 100 bp
136948 138948 contig of 100 bp
138948 140948 contig of 100 bp
140948 142948 contig of 100 bp
142948 144948 contig of 100 bp
144948 146948 contig of 100 bp
146948 148948 contig of 100 bp
148948 150948 contig of 100 bp
150948 152948 contig of 100 bp
152948 154948 contig of 100 bp
154948 156948 contig of 100 bp
156948 158948 contig of 100 bp
158948 160948 contig of 100 bp
160948 162948 contig of 100 bp
162948 164948 contig of 100 bp
164948 166948 contig of 100 bp
166948 168948 contig of 100 bp
168948 170948 contig of 100 bp
170948 172948 contig of 100 bp
172948 174948 contig of 100 bp
174948 176948 contig of 100 bp
176948 178948 contig of 100 bp
178948 180948 contig of 100 bp
180948 182948 contig of 100 bp
182948 184948 contig of 100 bp
184948 186948 contig of 100 bp
186948 188948 contig of 100 bp
188948 190948 contig of 100 bp
190948 192948 contig of 100 bp
192948 194948 contig of 100 bp
194948 196948 contig of 100 bp
196948 198948 contig of 100 bp
198948 200948 contig of 100 bp
200948 202948 contig of 100 bp
202948 204948 contig of 100 bp
204948 206948 contig of 100 bp
206948 208948 contig of 100 bp
208948 210948 contig of 100 bp
210948 212948 contig of 100 bp
212948 214948 contig of 100 bp
214948 216948 contig of 100 bp
216948 218948 contig of 100 bp
218948 220948 contig of 100 bp
220948 222948 contig of 100 bp
222948 224948 contig of 100 bp
224948 226948 contig of 100 bp
226948 228948 contig of 100 bp
228948 230948 contig of 100 bp
230948 232948 contig of 100 bp
232948 234948 contig of 100 bp
234948 236948 contig of 100 bp
236948 238948 contig of 100 bp
238948 240948 contig of 100 bp
240948 242948 contig of 100 bp
242948 244948 contig of 100 bp
244948 246948 contig of 100 bp
246948 248948 contig of 100 bp
248948 250948 contig of 100 bp
250948 252948 contig of 100 bp
252948 254948 contig of 100 bp
254948 256948 contig of 100 bp
256948 258948 contig of 100 bp
258948 260948 contig of 100 bp
260948 262948 contig of 100 bp
262948 264948 contig of 100 bp
264948 266948 contig of 100 bp
266948 268948 contig of 100 bp
268948 270948 contig of 100 bp
270948 272948 contig of 100 bp
272948 274948 contig of 100 bp
274948 276948 contig of 100 bp
276948 278948 contig of 100 bp
278948 280948 contig of 100 bp
280948 282948 contig of 100 bp
282948 284948 contig of 100 bp
284948 286948 contig of 100 bp
286948 288948 contig of 100 bp
288948 290948 contig of 100 bp
290948 292948 contig of 100 bp
292948 294948 contig of 100 bp
294948 296948 contig of 100 bp
296948 298948 contig of 100 bp
298948 300948 contig of 100 bp
300948 302948 contig of 100 bp
302948 304948 contig of 100 bp
304948 306948 contig of 100 bp
306948 308948 contig of 100 bp
308948 310948 contig of 100 bp
310948 312948 contig of 100 bp
312948 314948 contig of 100 bp
314948 316948 contig of 100 bp
316948 318948 contig of 100 bp
318948 320948 contig of 100 bp
320948 322948 contig of 100 bp
322948 324948 contig of 100 bp
324948 326948 contig of 100 bp
326948 328948 contig of 100 bp
328948 330948 contig of 100 bp
330948 332948 contig of 100 bp
332948 334948 contig of 100 bp
334948 336948 contig of 100 bp
336948 338948 contig of 100 bp
338948 340948 contig of 100 bp
340948 342948 contig of 100 bp
342948 344948 contig of 100 bp
344948 346948 contig of 100 bp
346948 348948 contig of 100 bp
348948 350948 contig of 100 bp
350948 352948 contig of 100 bp
352948 354948 contig of 100 bp
354948 356948 contig of 100 bp
356948 358948 contig of 100 bp
358948 360948 contig of 100 bp
360948 362948 contig of 100 bp
362948 364948 contig of 100 bp
364948 366948 contig of 100 bp
366948 368948 contig of 100 bp
368948 370948 contig of 100 bp
370948 372948 contig of 100 bp
372948 374948 contig of 100 bp
374948 376948 contig of 100 bp
376948 378948 contig of 100 bp
378948 380948 contig of 100 bp
380948 382948 contig of 100 bp
382948 384948 contig of 100 bp
384948 386948 contig of 100 bp
386948 388948 contig of 100 bp
388948 390948 contig of 100 bp
390948 392948 contig of 100 bp
392948 394948 contig of 100 bp
394948 396948 contig of 100 bp
396948 398948 contig of 100 bp
398948 400948 contig of 100 bp
400948 402948 contig of 100 bp
402948 404948 contig of 100 bp
404948 406948 contig of 100 bp
406948 408948 contig of 100 bp
408948 410948 contig of 100 bp
410948 412948 contig of 100 bp
412948 414948 contig of 100 bp
414948 416948 contig of 100 bp
416948 418948 contig of 100 bp
418948 420948 contig of 100 bp
420948 422948 contig of 100 bp
422948 424948 contig of 100 bp
424948 426948 contig of 100 bp
426948 428948 contig of 100 bp
428948 430948 contig of 100 bp
430948 432948 contig of 100 bp
432948 434948 contig of 100 bp
4349
```

```

misc_feature /note="assembly_fragment"
120947..122766
/note="assembly_fragment"
122867..124644
/note="assembly_fragment"
124745..126582
/note="assembly_fragment"
126683..128325
/note="assembly_fragment"
128426..129436
/note="assembly_fragment"
129537..130773
/note="assembly_fragment"
130874..132008
/note="assembly_fragment"
132109..133187
/note="assembly_fragment"
133288..134426
/note="assembly_fragment"
134527..135528
/note="assembly_fragment"
135629..136678
/note="assembly_fragment"

```

Query Match 100.0%; Score 22; DB 2; Length 136678;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCATGTACAGGTAGAAAAGCC 22
Db 70734 GCATGTACAGGTAGAAAAGCC 70755

```

## RESULT 12

```

AC013606 AC013606 143899 bp DNA linear HTG 12-MAR-2000
LOCUS Homo sapiens clone RP11-10E21, WORKING DRAFT SEQUENCE, 14 unordered
DEFINITION pieces.
ACCESSION AC013606
VERSION AC013606.5 GI:7230045
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 143899)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,B.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 143899)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Doneilan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission

```

```

TITLE Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 12, 2000 this sequence version replaced gi:6850474.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

```

```

Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3055
Center clone name: 10_E_21
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 138658 bases at least Q40
Consensus quality: 141075 bases at least Q30
Consensus quality: 141906 bases at least Q20
Insert size: 176000; agarose-fp
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 5.8 in Q20 bases; sum-of-contigs

```

\* NOTE: This is a 'working draft' sequence. It currently consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 2043: contig of 2043 bp in length
2044 2143: gap of 100 bp
2144 6624: contig of 4481 bp in length
6625 6724: gap of 100 bp
6725 14152: contig of 7428 bp in length
14153 14252: gap of 100 bp
14253 19511: contig of 5259 bp in length
19512 19611: gap of 100 bp
19612 22594: contig of 2983 bp in length
22595 27624: contig of 5130 bp in length
27625 27924: gap of 100 bp
27925 34952: contig of 7028 bp in length
34953 35052: gap of 100 bp
35053 40782: contig of 5730 bp in length
40783 40882: gap of 100 bp
40883 48999: contig of 8117 bp in length
49000 62883: contig of 13784 bp in length
62884 62984: gap of 100 bp
62985 75227: contig of 12544 bp in length
75228 75628: gap of 100 bp
75629 87943: contig of 12316 bp in length
87944 88043: gap of 100 bp
88044 111029: contig of 22986 bp in length
111030 111129: gap of 100 bp
111130 143899: contig of 32770 bp in length.

```

## FEATURES

```

source
1..143899
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-10E21"
/clone_lib="RPC1-11 Human Male BAC"
1..2043
/note="assembly_fragment"
2144..6624
/note="assembly_fragment"
6725..14152
/note="assembly_fragment"
14253..19511
/note="assembly_fragment"
clone_end:T7
vector_side:right
19612..22594
/note="assembly_fragment"
clone_end:SP6

```

```

vector side:left"
misc_feature 22695..27824
  /note="assembly_fragment"
misc_feature 27925..34952
  /note="assembly_fragment"
misc_feature 35053..40782
  /note="assembly_fragment"
misc_feature 40883..48999
  /note="assembly_fragment"
misc_feature 49100..62883
  /note="assembly_fragment"
misc_feature 62984..75527
  /note="assembly_fragment"
misc_feature 75628..87943
  /note="assembly_fragment"
misc_feature 88044..111029
  /note="assembly_fragment"
misc_feature 111130..143899
  /note="assembly_fragment"

ORIGIN
Query Match 100.0%; Score 22; DB 2; Length 143899;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22
Db 38486 GCATGTTACAGGTAGAAAAGCC 38507

RESULT 13
AX411543
LOCUS AX411543 143899 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 15 from Patent WO0229059.
ACCESSION AX411543
VERSION AX411543.1 GI:21444137
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
  AUTHORS Sah,D.W.Y., Cate,R.L. and Strittmatter,S.M.
  TITLE Nogo receptor homologs
  JOURNAL Patent: WO 0229059-A 15-APR-2002;
  BIOGEN INC (US)
FEATURES
  source
    location/Qualifiers
      1..143899
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 22; DB 6; Length 143899;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22
Db 38486 GCATGTTACAGGTAGAAAAGCC 38507

RESULT 14
AC087505/c
LOCUS AC087505 165434 bp DNA linear HTG 07-JUL-2001
DEFINITION Homo sapiens chromosome 11 clone RP11-676P16 map 11, WORKING DRAFT
SEQUENCE, 12 unordered pieces.
ACCESSION AC087505
VERSION AC087505.4 GI:14626336
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

---

```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 165434)
  Birren,B., Linton,L., Nuebaum,C. and Lander,E.
  Homo sapiens chromosome 11, clone RP11-676P16
  Unpublished
2 (bases 1 to 165434)
  Birren,B., Linton,L., Nuebaum,C., Lander,E., Allen,N., Anderson,S.,
  Barna,N., Baetien,V., Boguslavsky,L., Boukhgalter,B., Brown,A.,
  Canarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
  Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
  Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
  Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
  Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
  Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
  Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
  Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
  McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
  Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
  O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
  Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
  Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
  Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
  Sougnaz,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
  Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
  Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
  Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
  Zembek,L., Zimmer,A. and Zody,M.
  Direct Submission
Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 7, 2001 this sequence version replaced gi:14029937.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: 676_P_16
Center clone name: L12298
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 160902 bases at least Q40
Consensus quality: 163021 bases at least Q30
Consensus quality: 163793 bases at least Q20
Insert size: 164000; agarose-fp
Insert size: 164334; sum-of-contigs
Quality coverage: 12.9 in Q20 bases; agarose-fp
Quality coverage: 12.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 15621: contig of 15621 bp in length
* 15622 15721: gap of 100 bp
* 15722 16822: contig of 1101 bp in length
* 16822 16922: gap of 100 bp
* 16923 18786: contig of 1864 bp in length
* 18787 18886: gap of 100 bp
* 18887 20784: contig of 1898 bp in length
* 20785 20884: gap of 100 bp
* 20885 22630: contig of 1746 bp in length
* 22730 22731: gap of 100 bp
* 22731 25643: contig of 2913 bp in length

```

\* 25644 25743: gap of 100 bp  
 \* 25744 27851: contig of 2108 bp in length  
 \* 27852 27951: gap of 100 bp  
 \* 27952 108933: contig of 80882 bp in length  
 \* 108934 108933: gap of 100 bp  
 \* 108934 122115: contig of 13182 bp in length  
 \* 122116 122216: gap of 100 bp  
 \* 122216 145049: contig of 22834 bp in length  
 \* 145050 145149: gap of 100 bp  
 \* 145150 162807: contig of 17658 bp in length  
 \* 162808 162907: gap of 100 bp  
 \* 162908 165434: contig of 2527 bp in length.

## FEATURES

## source

1. .165434  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11"  
 /clone="RP11-676P16"  
 /clone\_lib="RPC1-11 Human Male BAC"

## misc\_feature

1. .15621  
 /note="assembly\_fragment  
 clone\_end:SP6  
 vector\_side:left"

## misc\_feature

15722..16822

## misc\_feature

/notes="assembly\_fragment"

## misc\_feature

16923..18786

## misc\_feature

/note="assembly\_fragment"

## misc\_feature

18887..20784

## misc\_feature

/notes="assembly\_fragment"

## misc\_feature

20885..22630

## misc\_feature

/notes="assembly\_fragment"

## misc\_feature

22731..25643

## misc\_feature

/note="assembly\_fragment"

## misc\_feature

25744..27851

## misc\_feature

/notes="assembly\_fragment"

## misc\_feature

27952..108833

## misc\_feature

/notes="assembly\_fragment"

## misc\_feature

108934..122115

## misc\_feature

/note="assembly\_fragment"

## misc\_feature

122216..145049

## misc\_feature

/notes="assembly\_fragment"

## misc\_feature

145150..162807

## misc\_feature

/note="assembly\_fragment"

## misc\_feature

162908..165434

## misc\_feature

/note="assembly\_fragment  
 clone\_end:T7  
 vector\_side:right"

## ORIGIN

Query Match 100.0%; Score 22; DB 2; Length 165434;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps -0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22

Db 144536 GCATGTTACAGGTAGAAAAGCC 144515

## RESULT 15

## AP002893

LOCUS AP002893 178000 bp DNA linear PRI 30-MAR-2002  
 DEFINITION Homo sapiens genomic DNA, chromosome 11q clone:RP11-624G17,  
 complete sequence.

## ACCESSION

## AP002893

VERSION AP002893.3 GI:19879819

## KEYWORDS

## HTG.

## SOURCE

## Homo sapiens

## ORGANISM

## Homo sapiens (human)

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

## 1

## AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
 Homo sapiens genomic DNA  
 Published Only in Database (2000)

## JOURNAL

## REFERENCE

## 2 (bases 1 to 178000)

## Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

## Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

## Direct Submission

## TITLE

## JOURNAL

## Submitted (16-OCT-2000) Masahira Hattori, The Institute of Physical

## and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

## 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

## (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,

## Tel:81-45-503-9111, Fax:81-45-503-9170)

## On Apr 1, 2002 this sequence version replaced gi:13469698.

## COMMENT

## FEATURES

## source

## 1. .178000

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## /chromosome="11"

## /map="11q"

## /clone="RP11-624G17"

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## 22; Conservative

## 0; Mismatches

## 0; Indels

## 0; Gaps

## 0;

## QY

## 1 GCATGTTACAGGTAGAAAAGCC 22

## Db

## 55502 GCATGTTACAGGTAGAAAAGCC 55523

## Search completed: April 10, 2004, 17:32:57

## Job time : 91.838 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 14:52:03 ; Search time 215.172 seconds  
(without alignments)  
5998.994 Million cell updates/sec

Title: US-09-743-825-1  
Perfect score: 2326  
Sequence: 1 ccgggctggagggggcaaa.....aggaagtggagaaaaaaa 2326

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/6C\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73.6	3.2	898	2	US-08-966-316-4
2	60	2.6	7218	1	US-08-232-463-14
3	43.2	1.9	1449	4	US-09-252-991A-12542
4	43.2	1.9	1452	4	US-09-252-991A-12542
5	42.8	1.8	2444	3	US-08-920-610-1
6	42.8	1.8	2444	4	US-09-023-655-1274
7	42.6	1.8	4403765	3	US-09-103-840A-2
8	42.6	1.8	4411529	3	US-09-103-840A-1
9	41.6	1.8	1376	4	US-09-016-434-1361
10	41.6	1.8	1394	2	US-08-068-729-3
11	41.6	1.8	1394	3	US-09-255-671-3
12	41.6	1.8	1394	4	US-09-395-366-3
13	41.2	1.8	526	4	US-09-072-596-290
14	41.2	1.8	526	4	US-09-072-967-295
15	40.2	1.7	1941	5	PCT-US91-07635-5
16	39.4	1.7	1965	3	US-09-178-252-26
17	39.4	1.7	1965	4	US-09-826-660-26
18	38.8	1.7	364	4	US-09-621-976-17202
19	38.6	1.7	1935	4	US-09-620-312D-236
20	37.8	1.6	843	4	US-09-252-991A-767
21	37.8	1.6	1278	4	US-09-252-991A-738
22	37.8	1.6	1359	4	US-09-252-991A-705
23	37.8	1.6	1587	4	US-09-252-991A-676
24	37.4	1.6	717	4	US-09-513-783A-47
25	37.4	1.6	1929	2	US-08-818-253-1
26	37.4	1.6	1929	3	US-08-818-253-5
27	37.4	1.6	1929	3	US-08-818-252-1

28	37.4	1.6	1929	3	US-08-818-252-5	Sequence 5, Appli
29	37.4	1.6	1959	2	US-08-818-253-3	Sequence 3, Appli
30	37.4	1.6	1959	3	US-08-818-252-3	Sequence 3, Appli
31	37.4	1.6	1971	2	US-08-818-253-7	Sequence 7, Appli
32	37.4	1.6	1971	3	US-08-818-252-7	Sequence 7, Appli
33	37.4	1.6	3171	4	US-09-513-783A-31	Sequence 31, Appli
34	37.4	1.6	4833	4	US-09-513-783A-21	Sequence 21, Appli
35	36.6	1.6	289	3	US-09-007-005-17	Sequence 17, Appli
36	36.6	1.6	289	3	US-09-244-796-17	Sequence 17, Appli
37	36.6	1.6	1344	4	US-09-252-991A-7741	Sequence 7741, Ap
38	36.6	1.6	2316	4	US-09-252-991A-7830	Sequence 7830, Ap
39	36.6	1.6	3144	4	US-09-252-991A-12554	Sequence 12554, A
40	36.6	1.6	4431	4	US-09-252-991A-12856	Sequence 12856, A
41	36.6	1.6	118067	4	US-09-497-855A-32	Sequence 32, Appli
42	36.4	1.6	3063	4	US-09-252-991A-12299	Sequence 12299, A
43	36.4	1.6	3105	4	US-09-252-991A-12350	Sequence 12350, A
44	36.2	1.6	948	1	US-08-154-915-3	Sequence 3, Appli
45	36.2	1.6	948	5	PCT-US93-09945-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-08-966-316-4  
; Sequence 4, Application US/08966316  
; Patent No. 5932445  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Reddy, Roopa  
; APPLICANT: Murty, Lynn E.  
; APPLICANT: Mathur, Preete  
; TITLE OF INVENTION: SIGNAL PEPTIDE - CONTAINING PROTEINS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/966,316  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0424 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 898 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BLADTUT06  
; CLONE: 1899132  
; US-08-966-316-4

```

Query Match      3.2%; Score 73.6; DB 2; Length 898;
Best Local Similarity 49.1%; Pred. No. 1.8e-10;
Matches 267; Conservative 0; Mismatches 259; Indels 18; Gaps 2;

QY 128 AGGGCTGTGCTGGAGAACTCTTCTCTGTGTACTCTGGGCTGGGGCTCCCTGTG 187
    |||||
Db 343 ACTGGCTGTGCTGGAATGCTGGCTTGTGGCGTCTTGTGGCTGGCTTCACTAGTG 402

QY 198 ATCATCTCTGAAGAAAGAGGGCTTCTATCCAGCAGTGGCCAGCTGAGAGCAGCAAC 247
    |||||
Db 403 TTGTGCTTCAAGAAATGAAGTACTTTAAGATCTGTGTGGACCAAGATCTGGGCGGATT 462

QY 248 ACCACCCAGGATGAGCAGCGCAGGTGGCCAGCTGTGACAGCAGCAGAGATGCTCAAC 307
    |||||
Db 463 GCGATGCCACAGGCA-----GGCTGACTGCAAGCCAGCAGATGAGGTTCTCA 513

QY 308 CTGGGCTTACCAATGGTTCCTTCTGTGTGCTCAGCGCCACACCTCCCTCCTGAGATCCTC 367
    |||||
Db 514 CTCATCTTCACTGCTGGGCTCTTCTATGAACAATTTCTATGACATTTCCCACTGGCTACATC 573

QY 368 ATGGACCGCTTTGGCCCCCGACCGTGGGCTGGTGGAGTGGCTGCTTCACTGCTGCC 427
    |||||
Db 574 TTGTGACCGTTCACAGACCAACCGTGGCAGCGCTCATAGCCATATTTTCTACACCAACCGCC 633

QY 428 TGCACCTCATGCGCTGGCCTCCCGGACGTGGAGCTCTGTCTCCGTTGATATTCCTG 487
    |||||
Db 634 ACACATCATAGACCTTCACTC-----TGACGGCTACGCGTGTCTTCTCTG 684

QY 488 GGGCTGTCCCTGAATGGCTTGGTGGCATTCGCTCAACGTTCACTTCACTCAGCTGCC 547
    |||||
Db 685 GCGATGCCAATGCTCACCATTGGGGAATCCTGTTTCTCATCACCACCTGCAGATTGGG 744

QY 548 ACATGTTTGGAACTGGCTCCAGCTTAAGTGGCCCTCATGATGGCTCTTACGCTCT 607
    |||||
Db 745 AACCTATTGGCAACACCGTTGACCATCATCTCTGTACAAATGGAGCATTTGACTCT 804

QY 608 TCTGCCATTACGTTCCAGGAATCAAGCTGATCTAGATGCGGCTGTGGCCTTCTGTGTC 667
    |||||
Db 805 TCTCGGAGTCTTCTTATTATTAGCTTCTTTATGAAGAAGCATCAGCTCAGGCCC 864

QY 668 ATCA 671
    ||
Db 865 TGCA 868

```

RESULT 2

```

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313

```

```

; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgt-F1s
; US-08-232-463-14

Query Match      2.6%; Score 60; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred. No. 3.4e-06;
Matches 18; Conservative 227; Mismatches 157; Indels 0; Gaps 0;

QY 329 TTCTGTCTCAGCGCCACACCCCTGCCACTGGGATCCTCATGAGCCGCTTTGGCCCCGA 388
    |||||
Db 1032 TCCGAGCTTGGCTGCGAGTGGAGGCTTGGATATTTTTTTTTTTTTTTTTTTTTT 1091

QY 389 CCCGTGGCGTGGTGGCAGTGCCTGTCTCACTGCTGCTGCACCTCATGCGCCCTGCGC 448
    |||||
Db 1092 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1151

QY 449 TCCGCGAGCTGGAAGCTGTCTCGTGTGATATCTCGGCGCTGCTCCTGAAATGGCTTT 508
    |||||
Db 1152 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1211

QY 509 GGTGGCATCTGCCTAACGTTTCACTCAGCTGCGCCACACATGTTGGGAACCTGCGC 568
    |||||
Db 1212 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1271

QY 569 TCCAGCTTAATGGCCCTCATGATGGCTCTTACGCTCTTCTGCAITTACGTTCCCGAG 628
    |||||
Db 1272 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1331

QY 629 ATCAAGCTGATCTAGATGCGGCTGCGCTTCTGCTGCTCATGCTTACCTGCTGCTGCG 688
    |||||
Db 1332 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1391

QY 689 CTGGCCTGCTTATCTTCTGAACCTGACCCCTCACTGGGCC 730
    |||||
Db 1392 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1433

RESULT 3
US-09-252-991A-12542
; Sequence 12542, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12542
; LENGTH: 1449
; TYPE: DNA

```





; Sequence 1274, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HERewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1274:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2444 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g307299  
; US-09-023-655-1274

Query Match 1.8%; Score 42.8; DB 4; Length 2444;  
Best Local Similarity 53.6%; Pred. No. 0.12;  
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
Qy 91 GCAACAGGCGTACCGAGCGCTGTGGATGGCCCTGACGGGTGTGCTGGAGAACCTCTT 150  
Db 2 GCACAGGCGGGCGGGGTTCGAGCTGGCGCGCGCATGGACGAATGTTCCCGCTCAT 61  
Qy 151 CTTCCTCTGCTGCT 210  
Db 62 CTTCCCGGAG 121  
Qy 211 TTTATTCAGCAGCTGCCAGCTGTAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 256  
Db 122 CTCGCGGGCGAGCATCCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 167

RESULT 7  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.

; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
; US-09-103-840A-2

Query Match 1.8%; Score 42.6; DB 3; Length 4403765;  
Best Local Similarity 49.8%; Pred. No. 7.5;  
Matches 108; Conservative 0; Mismatches 109; Indels 0; Gaps 0;  
Qy 663 TGGTCATCATGTTTCACTGCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 722  
Db 404171 TGGCGGCCAACTACACCGTGTGCACACACACCCAGCTGCTCAATCGTTGGTGGCGGACA 404112  
Qy 723 ACTGGCCCATCGAAGCCTTTCTCCCTCGCTGAGGAAGTCAATTACACGAAGAAGATCAAGC 782  
Db 404111 AGAGCGTGGTCCCTGTCACTCCGGTTTCTCAGGACATCCTTACACGACCCGCTGCTACC 404052  
Qy 783 TGAGTGGGCTGGCCCTGGACCAACAGGTGACAGTGACCTTCTTACACCATGTGACCA 842  
Db 404051 TGGGTTCGGCACACAAGGTCTACGAGGCACACACGGAGCTGATCGGTGCCGCGGGGCCA 403992  
Qy 843 CCATGGGCGAGGCTCAGCCAGAAAGGCCCGCCAGCCT 879  
Db 403991 GCCTGACGAGATGCCGCGGCATGCCGACCGCAGCTT 403955

RESULT 8  
US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
; US-09-103-840A-1

Query Match 1.8%; Score 42.6; DB 3; Length 4411529;  
Best Local Similarity 49.8%; Pred. No. 7.5; Indels 0; Gaps 0;  
Matches 108; Conservative 0; Mismatches 109; Indels 0; Gaps 0;  
Qy 663 TGGTCATCATGTTTCACTGCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 722  
Db 404116 TGGCGGCCAACTACACCGTGTGCACACACACCCAGCTGCTCAATCGTTGGTGGCGGACA 404057  
Qy 723 ACTGGCCCATCGAAGCCTTTCTCCCTCGCTGAGGAAGTCAATTACACGAAGAAGATCAAGC 782  
Db 404056 AGAGGCTGGTCCCTGTCACTCCGGTTTCTCAGGACATCCTTACACGACCCGCTGCTACC 403997

QY 783 TGAGTGGGTGGCCCTGGACCAAGGTGACAGGTGACCTTCTTACACCATGTGACCA 842  
Db 403996 TGGGTGGCACACAGGCTACGAGGACCAAGGCTGATCGGTGCGCGGGGGCCA 403937  
QY 843 CCATGGCCAGAGGCTCAGCCAGAAAGCCCGCCAGCCT 879  
Db 403936 GCCTGACCGAGATGCCGCGCCATCGCGACCGCAGCTT 403900

RESULT 9  
US-09-016-434-1361  
; Sequence 1361, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1361:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1376 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: 9410208  
US-09-016-434-1361

Query Match 1.8%; Score 41.6; DB 4; Length 1376;  
Best Local Similarity 50.0%; Pred. No. 0.18;  
Matches 104; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 302 CTCACCTGGGCTTACCAATTGGTTCCTTGGTCTCAGCGCCACACCTGCGCAGCTGGG 361  
Db 405 CACTTCTGGCGGCTGCAATGGTCTTCTGGGCTGTGCCGCTGCTGCTGGGCTGTGC 464  
QY 362 ATCCTCATGACCGCTTTGGCCCCCGACCGTGCAGTGGGAGTGCCTGCTTCACT 421  
Db 465 ATGGCGGTGAGCGCTGCGTGGCGGTGTCAGCGCGCGCTGCTCCACCGCGCGGCTCG 524  
QY 422 GCGTCTTGACCTCATGGCCCTGCGCTCCCGGAGCTGGAAGCTCTGTCTCGTTGATA 481  
Db 525 GTCCGCCGCGCGCTGGCGCTGGCGGTGGCGGGTGGCGGCTGGCGGCTGGCGGCTG 584

RESULT 11  
US-09-255-671-3  
; Sequence 3, Application US/09255671  
; Patent No. 6031079

QY 482 TTCTGGCGCTGTCCCTGAATGGCTTTG 509  
Db 585 CTGCCGCTGGCGGCTGGCGGCTATG 612  
RESULT 10  
US-08-068-729-3  
; Sequence 3, Application US/08068729  
; Patent No. 5985597  
; GENERAL INFORMATION:  
; APPLICANT: Ford-Hutchinson, Anthony  
; APPLICANT: Funk, Colin  
; APPLICANT: Grygorczyk, Richard  
; APPLICANT: Metters, Kathleen  
; TITLE OF INVENTION: DNA Encoding Prostaglandin Receptor EPI  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JOHN W. WALLEN III  
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.  
; CITY: RAHWAY  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/068,729  
; FILING DATE: 26-MAY-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WALLEN, JOHN W III  
; REGISTRATION NUMBER: 35,403  
; REFERENCE/DOCKET NUMBER: 19012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 594-3905  
; TELEFAX: (908) 594-4720  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1394 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-068-729-3

Query Match 1.8%; Score 41.6; DB 2; Length 1394;  
Best Local Similarity 50.0%; Pred. No. 0.18;  
Matches 104; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 302 CTCACCTGGGCTTACCAATTGGTTCCTTGGTCTCAGCGCCACACCTGCGCAGCTGGG 361  
Db 405 CACTTCTGGCGGCTGCAATGGTCTTCTGGGCTGTGCCGCTGCTGCTGGGCTGTGC 464  
QY 362 ATCCTCATGACCGCTTTGGCCCCCGACCGTGCAGTGGGAGTGCCTGCTTCACT 421  
Db 465 ATGGCGGTGAGCGCTGCGTGGCGGTGTCAGCGCGCGCTGCTCCACCGCGCGGCTCG 524  
QY 422 GCGTCTTGACCTCATGGCCCTGCGCTCCCGGAGCTGGAAGCTCTGTCTCGTTGATA 481  
Db 525 GTCCGCCGCGCGCTGGCGCTGGCGGTGGCGGGTGGCGGCTGGCGGCTGGCGGCTG 584  
QY 482 TTCTGGCGCTGTCCCTGAATGGCTTTG 509  
Db 585 CTGCCGCTGGCGGCTGGCGGCTATG 612

RESULT 11  
US-09-255-671-3  
; Sequence 3, Application US/09255671  
; Patent No. 6031079





```

; APPLICATION NUMBER: PCT/US91/07635
; FILING DATE: 19911018
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,929
; REFERENCE/DOCKET NUMBER: CRP-056PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1941 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS
; TISSUE TYPE: HIPPOCAMPUS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 507..1703
; OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
; OTHER INFORMATION: /product= "HOP2"
; OTHER INFORMATION:
;
PCT-US91-07635-5

Query Match 1.7%; Score 40.2; DB 5; Length 1941;
Best Local Similarity 49.3%; Pred. No. 0.54;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 334 GCTCAGCGCCACACCCCTGCCACTGGGGATCCTCATGACCGCTTTGGCCCCCGACCCCGT 393
Db |||||
QY 486 GCTGAGCGCCCGGCGCTGCCATGACCGCGCTCCCGCGCGCTCTGGCTCTGGGGCTG 545
Db |||||
QY 394 GCGGCTGTTGGCAGTGCCTTCTCCTGCGCTCTGACCTGACCCCTCATGGCCCTGCCCG 453
Db |||||
QY 546 GCGCTATGCGCGTGGCGGGGGCGGCCCGCGCTCGGACCCCGCGGCTGTCCCCAG 605
Db |||||
QY 454 GGACGTGGAAGCTCTGTCTCCGTTGATATTCCTGCGCTGTCCCTCAATGGCTTTGGTGG 513
Db |||||
QY 606 CGACGCTCTGGCGCGCGCGACCCGGGACGTGACGCGGAGATCCTGGCGGTGCTCGGGCTG 665
Db |||||
QY 514 CATCTGCCTAAGTTCACTTCACTTCACTACGCTGCC 546
Db |||||
QY 666 CCTGGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 698
Db |||||
```

Search completed: April 10, 2004, 20:25:38  
Job time : 228.172 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:11:52 ; Search time 9392.6 Seconds  
(without alignments)  
10733.545 Million cell updates/sec

Title: US-09-743-825-1  
Perfect score: 2326  
Sequence: 1 ccggggctgaggggggcaaa.....agggagtgagaaaaaaa 2326

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_hg:\*

3: gb\_in:\*

4: gb\_ov:\*

5: gb\_pat:\*

6: gb\_ph:\*

7: gb\_pl:\*

8: gb\_pr:\*

9: gb\_ro:\*

10: gb\_sy:\*

11: gb\_un:\*

12: gb\_vt:\*

13: gb\_wt:\*

14: gb\_xl:\*

15: gb\_yt:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_or:\*

21: em\_ov:\*

22: em\_pat:\*

23: em\_ph:\*

24: em\_pi:\*

25: em\_ro:\*

26: em\_se:\*

27: em\_un:\*

28: em\_vt:\*

29: em\_wt:\*

30: em\_xl:\*

31: em\_yt:\*

32: em\_hg\_inv:\*

33: em\_hg\_mus:\*

34: em\_hg\_pln:\*

35: em\_hg\_rod:\*

36: em\_hg\_mam:\*

37: em\_hg\_vrt:\*

38: em\_sv:\*

39: em\_higo\_hum:\*

40: em\_higo\_mus:\*

41: em\_higo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2326	100.0	2326	9	AF045584 Homo sapi
2	2299.2	98.8	2375	9	BC001639 Homo sapi
3	2234.6	96.1	2525	9	AB103033 Homo sapi
4	2214.6	95.2	2373	6	BD063631 Human pro
5	2128.4	91.5	2233	6	AX768966 Sequence
6	2066.6	88.8	2093	6	AX880371 Sequence
7	2066.6	88.8	2093	6	BD158342 Primer fo
8	2066.6	88.8	2093	9	AK027732 Homo sapi
9	1677	72.1	1677	6	BD063606 Human pro
10	1318	56.7	2507	10	BC053747 Mus muscu
11	684.2	29.4	136678	2	AP001332 Homo sapi
12	684.2	29.4	143899	2	AC013606 Homo sapi
13	684.2	29.4	143899	6	AX411543 Sequence
14	684.2	29.4	187566	2	AC015865 Homo sapi
15	682.6	29.3	165434	2	AC087505 Homo sapi
16	682.6	29.3	178000	9	AP002893 Homo sapi
17	627.4	27.0	632	6	AX869025 Sequence
18	627.4	27.0	632	6	BD149087 Primer fo
19	614.6	26.4	3024	9	BC027923 Homo sapi
20	563.6	24.2	4389	10	BC042513 Mus muscu
21	545.8	23.5	2899	5	BC044509 Danio rer
22	540	23.2	2643	6	BD127551 Primer fo
23	540	23.2	2643	9	AK075173 Homo sapi
24	503	21.6	2589	6	BD127459 Primer fo
25	503	21.6	2589	9	AK075068 Homo sapi
26	492.2	21.2	579	6	AX874028 Sequence
27	492.2	21.2	579	6	BD154090 Primer fo
28	423.6	18.2	464	6	AX332309 Sequence
29	423.6	18.2	464	6	AX407892 Sequence
30	376.8	16.2	2349	6	BD127568 Primer fo
31	376.8	16.2	2349	9	AK075196 Homo sapi
32	370.8	15.9	2445	9	BC019562 Homo sapi
33	289	12.4	630	6	BD125250 Primer fo
34	289	12.4	630	6	BD126512 Primer fo
35	277.4	11.9	1359	6	BD210015 Human tra
36	271.6	11.7	457	6	AX407361 Sequence
37	263.6	11.3	321	11	G22380 human STS W
38	262.6	11.3	412	11	G21539 human STS W
39	252.6	10.9	613	6	BD125149 Primer fo
40	252.6	10.9	613	6	BD126413 Primer fo
41	194.2	8.3	800	6	BD125271 Primer fo
42	194.2	8.3	800	6	BD126531 Primer fo
43	181.8	7.8	143899	2	AC013606 Homo sapi
44	181.8	7.8	143899	6	AX411543 Sequence
45	177.4	7.6	214711	2	AC108295 Rattus no

ALIGNMENTS

RESULT 1  
AF045584  
LOCUS Homo sapiens PB39 mRNA, complete cds.  
DEFINITION Homo sapiens PB39 mRNA, complete cds.  
ACCESSION AF045584  
VERSION AF045584.1 GI:3462514  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2326)  
AUTHORS Cole, K.A., Chuaqui, R.F., Katz, K., Pack, S., Zhuang, Z., Cole, C.E.,  
Lyne, J.C., Linehan, W.M., Liotta, L.A. and Emmert-Buck, M.R.  
TITLE cDNA sequencing and analysis of POV1 (PB39): a novel gene

JOURNAL	up-regulated in prostate cancer	Db	481	ATTCTGCGGCTGTCCTGTAATGGCTTTGGTGGCATCTGCCTAAAGTTTCACTTCACTCAC	540
MEDLINE	Genomics 51 (2), 282-287 (1998)	Qy	541	GCTGCCCAACATGTTTGGGAACCTTGGCTTCAAGTTTAAATGGCCCTCATGATTGGCTCTTA	600
PUBMED	98390192	Db	541	GCTGCCCAACATGTTTGGGAACCTTGGCTTCAAGTTTAAATGGCCCTCATGATTGGCTCTTA	600
REFERENCE	2 (bases 1 to 2326)	Qy	601	CGCTCTTCTGCCAATTAGTTTCCAGGAATCAAGCTGATCTACGATGCGGTGGCCCTT	660
AUTHORS	Cole, K.A., Chuquai, R., Katz, K., Pack, S., Zhuang, Z., Cole, C.E., Lyne, J.C., Linehan, M., Liotta, L.A. and Emmert-Buck, M.R.	Db	601	CGCTCTTCTGCCAATTAGTTTCCAGGAATCAAGCTGATCTACGATGCGGTGGCCCTT	660
TITLE	Direct Submission	Qy	661	CGTGGTCATCATGTTACCTGCTGCGCTGGCCCTGCTTATCTTTCTGAACATGCACCT	720
JOURNAL	Submitted (31-JAN-1998) Laboratory of Pathology, NCI/NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA	Db	661	CGTGGTCATCATGTTACCTGCTGCGCTGGCCCTGCTTATCTTTCTGAACATGCACCT	720
FEATURES	Location/Qualifiers	Qy	721	CAACTGGCCCATCGAAGCTTTCTGCCCCCTGAGGAAGTCAATTACAGGAAGATCAA	780
source	1..2326	Db	721	CAACTGGCCCATCGAAGCTTTCTGCCCCCTGAGGAAGTCAATTACAGGAAGATCAA	780
	/organism="Homo sapiens"	Qy	781	GCTGAGTGGGCTGGCCCTGGACCAAGAGTGCAGAGTGCACCTTCTTACACCCATGTGAC	840
	/mol_type="mRNA"	Db	781	GCTGAGTGGGCTGGCCCTGGACCAAGAGTGCAGAGTGCACCTTCTTACACCCATGTGAC	840
	/db_xref="taxon:9606"	Qy	841	CACCATGGCCAGAGGCTCAGCCAGAGGCCCCCAGCTGGAGGACGTTCCGGATGCCCT	900
	/chromosomes="11"	Db	841	CACCATGGCCAGAGGCTCAGCCAGAGGCCCCCAGCTGGAGGACGTTCCGGATGCCCT	900
	/map="11p11.1-p11.2"	Qy	901	CATGTCACCCAGAGATGTTCCGGGACCTCAGAAAACTTCTCTGAGAGGCTCTGCCCTT	960
	77..1756	Db	901	CATGTCACCCAGAGATGTTCCGGGACCTCAGAAAACTTCTCTGAGAGGCTCTGCCCTT	960
	/note="up-regulated in prostate cancer"	Qy	961	ACGCAAGAGCTCTGCTCCCCACCTTCTCTGAGAGCTCTCTCACCACCTGGGACGACCA	1020
	/codon_start=1	Db	961	ACGCAAGAGCTCTGCTCCCCACCTTCTCTGAGAGCTCTCTCACCACCTGGGACGACCA	1020
	/product="PB39"	Qy	1021	GCTCGGATCATCTTCTACATGGCTGTGTGAACAAGATGCTGGAGTACCTTGTGACTGG	1080
	/db_xref="GI:3462515"	Db	1021	GCTCGGATCATCTTCTACATGGCTGTGTGAACAAGATGCTGGAGTACCTTGTGACTGG	1080
	/translation="NAPLQOYRRRMMWMACTAVLENLFFSAVLIGWGLIILKNEG FYSSCPAESNTTQDEORWPGDQDDMLNLGFTIGSFVLSATLPLGLIMDRFG PRPVLVGSACFTASCTLMALASRDVEALSPILFIALSINGFGGICLFTSLTLPNMF GNLRSLMALMTGSAASSNAIPFGIKLIYDAGVAFVIMFTWSGLACILPINCILNWP IEAFPAPENVTKIKLUSGALDHKVTGDLFYTHVTMQRSLQKAPSLSDSDAFM SPQDVHGTSENPERSVPLRSLCPTFLMSLLTNGMTQLRIIFYMAAVNKLELYVT GGOEHETNEQQQVATQVTFYSSVFCAQMLLCLTCTPLIGYIMWRIRKXDCVDAPTQGT VLGDARDGVAIKSIRPRYCKIOKLTNAISAPLTLLLVGFGITCLINLHLQFTFV LHTIVRPFPHSAGSLYAAVPSNHFGLTGLQSLI SAVFALLQOPLFAMVUGPLKGE PFWNLGLLLFLSLFLPSLYFYRRLQBEYANGMPLKVLGSEVTA"	Qy	1081	TGGCCAGGAGCATGAGCAAAATGAAACAAACAAAGGTGGCAGAGACAGTTGGTTCTA	1140
		Db	1081	TGGCCAGGAGCATGAGCAAAATGAAACAAACAAAGGTGGCAGAGACAGTTGGTTCTA	1140
		Qy	1141	CTCTCTCGTCTTCCGGGCGCATGACAGTGTGTGCTTCTCACCCTGCCCTCATTTGGCTA	1200
		Db	1141	CTCTCTCGTCTTCCGGGCGCATGACAGTGTGTGCTTCTCACCCTGCCCTCATTTGGCTA	1200
		Qy	1201	CATCATGGAATGCGCGGATCAAGGACTGCGTGGAGCGCCCAACTCAGGACATGCTCGG	1260
		Db	1201	CATCATGGAATGCGCGGATCAAGGACTGCGTGGAGCGCCCAACTCAGGACATGCTCGG	1260
		Qy	1261	AGATGCCAGGAGCGGGGTTGCTACCAATCCATCAGACACGCTACTGCAAGATCCAAA	1320
		Db	1261	AGATGCCAGGAGCGGGGTTGCTACCAATCCATCAGACACGCTACTGCAAGATCCAAA	1320
		Qy	1321	GCTCAACAATGCCATCAGTGCCTTCACTGACCAACTCTGCTGTGTGGGTTTGGCAT	1380
		Db	1321	GCTCAACAATGCCATCAGTGCCTTCACTGACCAACTCTGCTGTGTGGGTTTGGCAT	1380
		Qy	1381	CACCTGTCTCATCAACAACCTTACACCTCCAGTTGTGACCTTTGCTGTCACACCATGT	1440
		Db	1381	CACCTGTCTCATCAACAACCTTACACCTCCAGTTGTGACCTTTGCTGTCACACCATGT	1440
		Qy	1441	TCGAGGTTTCTTCCACTCAGGCTGTGGAGTCTCTATGCTGAGTGTCCCATCCAAACA	1500
		Db	1441	TCGAGGTTTCTTCCACTCAGGCTGTGGAGTCTCTATGCTGAGTGTCCCATCCAAACA	1500
		Qy	1501	CTTTGGGACGCTGACAGGCTGTGACCTCCTCATCAGTGTGTGTTCGCTTGTCTGACGA	1560
		Db	1501	CTTTGGGACGCTGACAGGCTGTGACCTCCTCATCAGTGTGTGTTCGCTTGTCTGACGA	1560
		Qy	1561	GCACCTTTTATGCGGATGTGGGACCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGG	1620
		Db	1561	GCACCTTTTATGCGGATGTGGGACCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGG	1620

ORIGIN

Query Match	100.0%; Score 2326; DB 9; Length 2326;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 2326; Conservative	0; Mismatch 0; Indels 0; Gaps 0;	
Qy	1	CGGGGCTGGAGGGGGGCAAGCGGTTCCGAGGTGCAAGCTGTGCTGCCCGAGCCCTGC 60
Db	1	CGGGGCTGGAGGGGGGCAAGCGGTTCCGAGGTGCAAGCTGTGCTGCCCGAGCCCTGC 60
Qy	61	GGAGCTCGGGGCGAGCATGGCCCCCAGCGTCAACAGGCGTACCGAGGCGCTGTGGAT 120
Db	61	GGAGCTCGGGGCGAGCATGGCCCCCAGCGTCAACAGGCGTACCGAGGCGCTGTGGAT 120
Qy	121	GGCTGCAAGCTGTGCTGGAGAACTTCTTCTCTGCTGTACTCTCTGGGCTGGGCTC 180
Db	121	GGCTGCAAGCTGTGCTGGAGAACTTCTTCTCTGCTGTACTCTCTGGGCTGGGCTC 180
Qy	181	CCGTGTCATCTGAGAGAGAGGGCTTCTATTTCCAGCAGTGCCTCCAGCTGAGAGCAG 240
Db	181	CCGTGTCATCTGAGAGAGAGGGCTTCTATTTCCAGCAGTGCCTCCAGCTGAGAGCAG 240
Qy	241	CACCAACACACCCAGGATGACGAGCGAGTGGCGAGCTGTGACAGCAGGAGAGAT 300
Db	241	CACCAACACACCCAGGATGACGAGCGAGTGGCGAGCTGTGACAGCAGGAGAGAT 300
Qy	301	GCTCAACCTGGCTTCAACATTTGTTCTTCTGCTGTCAGCGCCACACCTTGCACCTGGG 360
Db	301	GCTCAACCTGGCTTCAACATTTGTTCTTCTGCTGTCAGCGCCACACCTTGCACCTGGG 360
Qy	361	GATCTCATGAGCCGCTTGGCCCCCGAGCCGCTGGGCTGGTGGAGTGCCTGCTTAC 420
Db	361	GATCTCATGAGCCGCTTGGCCCCCGAGCCGCTGGGCTGGTGGAGTGCCTGCTTAC 420
Qy	421	TGCGTCTGACCTCATGAGCCGCTTGGCCCCCGAGCCGCTGGGCTGGTGGAGTGCCTGCTTAC 480
Db	421	TGCGTCTGACCTCATGAGCCGCTTGGCCCCCGAGCCGCTGGGCTGGTGGAGTGCCTGCTTAC 480
Qy	481	ATTCTGCGGCTGTCCCTGAATGGCTTGGTGGCATCTGCCCTAACGTTCACTTCACTCAC 540



```
QY 1621 CTTCTGCTATTCTCACTCTGGGATTCTGTGTGCTTCTTACCTCTTCTATTACCGTGC 1680
|||||
Db 1621 CTTCTGCTATTCTCACTCTGGGATTCTGTGTGCTTCTTACCTCTTCTATTACCGTGC 1680
|||||
QY 1681 CCGGCTCCAGCAGGAGTACGCGGCCAATGGATGGGCCCACTGAAGGTCTTAGCGGCTC 1740
|||||
Db 1681 CCGGCTCCAGCAGGAGTACGCGGCCAATGGATGGGCCCACTGAAGGTCTTAGCGGCTC 1740
|||||
QY 1741 TGAGGTGACCGCATAGACTTCTCAGACCAAGGAGGAGCTGATGACAGGCAATCAAGCCCTG 1800
|||||
Db 1741 TGAGGTGACCGCATAGACTTCTCAGACCAAGGAGGAGCTGATGACAGGCAATCAAGCCCTG 1800
|||||
QY 1801 AGCAACCAAAAGAGTGGCCCATATGGCTTTTCTACCTGTAAACATGCACATAGAGCCATG 1860
|||||
Db 1801 AGCAACCAAAAGAGTGGCCCATATGGCTTTTCTACCTGTAAACATGCACATAGAGCCATG 1860
|||||
QY 1861 GCGGTAGATTTAAATACAAAGAGAAAGTCTATTTTGTAAAGACTGCAAAAGAGGA 1920
|||||
Db 1861 GCGGTAGATTTAAATACAAAGAGAAAGTCTATTTTGTAAAGACTGCAAAAGAGGA 1920
|||||
QY 1921 AAAAAACCTTCAAAAACGCCCCCTAAGTCAAGCTCCATTGACTGAAGACAGTCCCTAT 1980
|||||
Db 1921 AAAAAACCTTCAAAAACGCCCCCTAAGTCAAGCTCCATTGACTGAAGACAGTCCCTAT 1980
|||||
QY 1981 CTTAGAGGGTGTGAGCTTTCTTCTCTCTTGGGTTGGAGGAGACAGGGTCCCTCTTATCT 2040
|||||
Db 1981 CTTAGAGGGTGTGAGCTTTCTTCTCTCTTGGGTTGGAGGAGACAGGGTCCCTCTTATCT 2040
|||||
QY 2041 CTTTCTAGCGGTGCTGCTCTCTGCTGCTCTTGGGGGATCGGCAACAGGCTACCCCTGA 2100
|||||
Db 2041 CTTTCTAGCGGTGCTGCTCTCTGCTGCTCTTGGGGGATCGGCAACAGGCTACCCCTGA 2100
|||||
QY 2101 GGTCCCATCTGCATGAGTGTGCACAAATGCAATGCTCTGTGTATGTGTAATGTGAG 2160
|||||
Db 2101 GGTCCCATCTGCATGAGTGTGCACAAATGCAATGCTCTGTGTATGTGTAATGTGAG 2160
|||||
QY 2161 AAAAAACAGCCCTCTTTTTCAGAAAGAAAGGGGCCCTGAGGTGCGACGTCTGTCTGGGTT 2220
|||||
Db 2161 AAAAAACAGCCCTCTTTTTCAGAAAGAAAGGGGCCCTGAGGTGCGACGTCTGTCTGGGTT 2220
|||||
QY 2221 AGGGGTGGGGTTCGCCCCCTTCCAGGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
|||||
Db 2221 AGGGGTGGGGTTCGCCCCCTTCCAGGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
|||||
QY 2281 GCTTGAAGTCTTAGAGGAATAAAAAAGGAGTGAAGAAAAA 2326
|||||
Db 2281 GCTTGAAGTCTTAGAGGAATAAAAAAGGAGTGAAGAAAAA 2326
|||||
```

```
RESULT 2
BC001639
LOCUS
DEFINITION
MGC:1944 IMAGE:2959372), complete cds.
ACCESSION
BC001639
VERSION
BC001639.2 GI:33876352
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2375)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Schetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
Carninci P., Prange C., Raja S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
```

Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalek U., Smalish D.E., Schnerch A., Schein J.E., Jones S.J. and Marra M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2375)

Strausberg R.

Direct Submission

Submitted (08-JAN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 19, 2003 this sequence version replaced gi:12804464.

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: [amadnan@systemsbiology.org](mailto:amadnan@systemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAL Plate: 2 Row: c Column: 12

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505970.

Location/Qualifiers

1..2375

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:1944 IMAGE:2959372"

/tissue\_type="Colon, adenocarcinoma"

/clone\_lib="NIH MGC\_15"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

1..2375

/gene="SLC43A1"

/note="synonyms: PB39, R00504, POVI"

/db\_xref="LocusID:8501"

/db\_xref="MIM:603733"

119..1798

/codon\_start=1

/product="solute carrier family 43, member 1"

/protein\_id="AAH01639.1"

/db\_xref="GI:12804465"

/db\_xref="LocusID:8501"

/translation="MAPTLQYRRRRRRMACTAVLENLFFSAVLGWSLLILKNEG FYSCTPAESSTNTTQDRMPGCDQDMLNGFTTGSFVLSATLPLGILMDRFG PRPVLVGSACFTASCTLMALASRDVEALSPLIFLALNGFGGICLTFTSLTNMF GNLRSTLMALMGVSASSAIPTPGIKLYDAGVAFVIMFTWSGLACILFNCTLNMP IEAPPEEVNVTYKKIKLSGLADHKVTGDLFYTHVTMTGQRLSKAPSLDGSDAEM SQDVRGTSENLPEPSVRLKSCSPTEFLWLLTGMQLRIIFYMAAVKNMLEYLT GVQHEHTEQQQKVAETGVFSSVFGAMQLCLLTPLIGYIMDRKDCVDAQVFTV VLGDAGVATKSRIPRYCKTQKLTNAISAFTLTNLLVFGITCLINLNHLPQFTFV LHTIVRGFHSACSLVAAPSPSHFTLTGLQSLISAVFALLQQLPFLFMAVYGLPKGE PFWNLGLLLFSLGLFLPLPSYLYFYRRLRQEQYAAANGMGPLKVLGSEVTA"

CDS

gene

FEATURES

source

ORIGIN

Query Match 98.8%; Score 2399.2; DB 9; Length 2375;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 2316; Conservative		0; Mismatches		8; Indels		2; Gaps		1;	
Qy	1	CGGGGCTGGAGGGGGGCAAGCGGTTCCGAGGTGCAAGCTGTGTCGCCGAGCCCTGC	60						
Db	43	CGGGGCTGGAGGGGGGCAAGCGGTTCCGAGGTGCAAGCTGTGTCGCCGAGCCCTGC	102						
Qy	61	GGAGCTCGGGGGCAGCATGGCCCCCAACGCTGCAAGCGGTACCGAGGCGCTGTGTGAT	120						
Db	103	GGAGCTCGGGGGCAGCATGGCCCCCAACGCTGCAAGCGGTACCGAGGCGCTGTGTGAT	162						
Qy	121	GGCTGCAAGCTGTGCTGGAGAACCTTCTTCTCTGCTGTAATCTCTGGGCTGGGCTC	180						
Db	163	GGCTGCAAGCTGTGCTGGAGAACCTTCTTCTCTGCTGTAATCTCTGGGCTGGGCTC	222						
Qy	181	CTGTGTGATCATCTGAGAACGAGGGCTTCTATTCCAGCAGCTGCCAGCTGAGAGCAG	240						
Db	223	CTGTGTGATCATCTGAGAACGAGGGCTTCTATTCCAGCAGCTGCCAGCTGAGAGCAG	282						
Qy	241	CACCAACACACCCAGGATGAGCAGCGAGGTGGCCAGGCTGTGACACGAGGACGAGAT	300						
Db	283	CACCAACACACCCAGGATGAGCAGCGAGGTGGCCAGGCTGTGACACGAGGACGAGAT	342						
Qy	301	GCTCAACCTGGGCTTCAACATTTGGTTCCTTCTGCTGCAAGCCACCAACCTGCCACTGG	360						
Db	343	GCTCAACCTGGGCTTCAACATTTGGTTCCTTCTGCTGCAAGCCACCAACCTGCCACTGG	402						
Qy	361	GATCTCATGAGCGCTTTGGCCCCCGGACCGCTGGGCTGGTGGAGCTGCTGCTTAC	420						
Db	403	GATCTCATGAGCGCTTTGGCCCCCGGACCGCTGGGCTGGTGGAGCTGCTGCTTAC	462						
Qy	421	TGCGTCTCTGACCTCATGGCCCTGGCCCTCCCGGACGCTGGAAGCTCTGCTCCGCTGAT	480						
Db	463	TGCGTCTCTGACCTCATGGCCCTGGCCCTCCCGGACGCTGGAAGCTCTGCTCCGCTGAT	522						
Qy	481	ATTCTCTGGGCTGTGCTGGAATGGCTTTGGTGGCATCTGCTTAAAGCTTCACTTCACT	540						
Db	523	ATTCTCTGGGCTGTGCTGGAATGGCTTTGGTGGCATCTGCTTAAAGCTTCACTTCACT	582						
Qy	541	GCTGCCCAACATGTTGGAACTCGGCTCCAGCTTAAAGCTTAAAGCTTAAAGCTTAA	600						
Db	583	GCTGCCCAACATGTTGGAACTCGGCTCCAGCTTAAAGCTTAAAGCTTAAAGCTTAA	642						
Qy	601	CGCTCTTCTGCAATTTACCTTCCAGGAATCAAGCTGATCTACGATGCGGCTGGCCCTT	660						
Db	643	CGCTCTTCTGCAATTTACCTTCCAGGAATCAAGCTGATCTACGATGCGGCTGGCCCTT	702						
Qy	661	CGTGTCTATCATGTTCACTGTCTGGCTGGCTGCTGCTTATCTTTCTGAACTGCAACCT	720						
Db	703	CGTGTCTATCATGTTCACTGTCTGGCTGGCTGCTTATCTTTCTGAACTGCAACCT	762						
Qy	721	CAACTGGCCCATCGAAGCCTTTCCTGCCCTGAGGAAGTCAATTACAGAGAGATCAA	780						
Db	763	CAACTGGCCCATCGAAGCCTTTCCTGCCCTGAGGAAGTCAATTACAGAGAGATCAA	822						
Qy	781	GCTGAGTGGGCTGGCCCTGGACCAAGGTGACAGGTGACCTTCTTACACCCATGTGAC	840						
Db	823	GCTGAGTGGGCTGGCCCTGGACCAAGGTGACAGGTGACCTTCTTACACCCATGTGAC	882						
Qy	841	CACCATGGGCCAGAGGCTCAGCCAGAGGCCCCAGCTGGAGGACGTTGGATGGCTT	900						
Db	883	CACCATGGGCCAGAGGCTCAGCCAGAGGCCCCAGCTGGAGGACGTTGGATGGCTT	942						
Qy	901	CATGTCAACCCAGGATGTTGGGGGACCTTCAGAAACCTTCTGAGAGGTCTGTCCCTT	960						
Db	943	CATGTCAACCCAGGATGTTGGGGGACCTTCAGAAACCTTCTGAGAGGTCTGTCCCTT	1002						
Qy	961	ACGCAAGAGCCTCTGCTCCCCCACTTCTCTGTGGAGCTTCTTCAACATGGGATGACCCA	1020						
Db	1003	ACGCAAGAGCCTCTGCTCCCCCACTTCTCTGTGGAGCTTCTTCAACATGGGATGACCCA	1062						
Qy	1021	GCTGGGATCATCTTCTATCATGGCTGTGTGAACAGATGTCTGTGTGTGTGTGTGTGTGT	1080						
Db	1063	GCTGGGATCATCTTCTATCATGGCTGTGTGAACAGATGTCTGTGTGTGTGTGTGTGTGT	1122						

Qy	1081	TGGCCAGGAGCATGAGACAAATGAACAGCAAAAAGTGGCAGAGACAGTTGGTTCTTA	1140
Db	1123	TGGCCAGGAGCATGAGACAAATGAACAGCAAAAAGTGGCAGAGACAGTTGGTTCTTA	1182
Qy	1141	CTCTCCGTCTTTCCGGGGCCATGACAGCTGTGTGCTTCTTCACTGCCCCCTCATTTGGCTA	1200
Db	1183	CTCTCCGTCTTTCCGGGGCCATGACAGCTGTGTGCTTCTTCACTGCCCCCTCATTTGGCTA	1242
Qy	1201	CATCATGGAATGCGGGGATCAAGGACTGCGTGGAGCGCCCCAACTCAGGGCAGTCTCTCGG	1260
Db	1243	CATCATGGAATGCGGGGATCAAGGACTGCGTGGAGCGCCCCAACTCAGGGCAGTCTCTCGG	1302
Qy	1261	AGATGCCAGGAGCGGGGTTGCTACCAATCCATCAGACCAAGCTTCTGAGAGTCCAAAA	1320
Db	1303	AGATGCCAGGAGCGGGGTTGCTACCAATCCATCAGACCAAGCTTCTGAGAGTCCAAAA	1362
Qy	1321	GCTCACCAGTCCATCAGTGCCTTCACTGACCAAACTGCTGCTGTGTGGGTTTGGCAT	1380
Db	1363	GCTCACCAGTCCATCAGTGCCTTCACTGACCAAACTGCTGCTGTGTGGGTTTGGCAT	1422
Qy	1381	CACCTGTCTCATCAACAACTTACACCTCCAGTTTGTGACCTTGTCTGTGACACATTTGT	1440
Db	1423	CACCTGTCTCATCAACAACTTACACCTCCAGTTTGTGACCTTGTCTGTGACACATTTGT	1482
Qy	1441	TCGAGGTTTCTTCCACTCAGCTGCGGAGTCTCTATGCTGACGTTTCCCATCCAAACA	1500
Db	1483	TCGAGGTTTCTTCCACTCAGCTGCGGAGTCTCTATGCTGACGTTTCCCATCCAAACA	1542
Qy	1501	CTTTGGGACGCTGACAGGCTTGCATCAGTGTCTGCTGTGTGCTGCTTGCCTTCAACA	1560
Db	1543	CTTTGGGACGCTGACAGGCTTGCATCAGTGTCTGCTGTGTGCTGCTTGCCTTCAACA	1602
Qy	1561	GCACCTTTTATGCGGATGTTGGGACCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGG	1620
Db	1603	GCACCTTTTATGCGGATGTTGGGACCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGG	1662
Qy	1621	CCTCTGCTATTTCTCACTTCTGCTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCT	1680
Db	1663	CCTCTGCTATTTCTCACTTCTGCTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCT	1722
Qy	1681	CCGGCTCCAGCAGGAGTACGCGCCAAATGGGATGGGCCCCCACTGAAGGTGCTTACGGCTC	1740
Db	1723	CCGGCTCCAGCAGGAGTACGCGCCAAATGGGATGGGCCCCCACTGAAGGTGCTTACGGCTC	1782
Qy	1741	TGAGGTGACCGCATAGACTTCTGACCAAGGAGCTTGGATGACAGGCAATCAAGGCTG	1800
Db	1783	TGAGGTGACCGCATAGACTTCTGACCAAGGAGCTTGGATGACAGGCAATCAAGGCTG	1842
Qy	1801	AGCAACCAAAAGGAGTGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1860
Db	1843	AGCAACCAAAAGGAGTGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1902
Qy	1861	GCCTAGATTTTATAAATCAAGAGAGTCTATTTTGTAAAGACTGCAAAAAAGGAGGA	1920
Db	1903	GCCTAGATTTTATAAATCAAGAGAGTCTATTTTGTAAAGACTGCAAAAAAGGAGGA	1962
Qy	1921	AAAAAAACCTTCAAAAAAGCCCCCTTAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1980
Db	1963	AAAAAAACCTTCAAAAAAGCCCCCTTAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2022
Qy	1981	CCTAGAGGGGTTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2040
Db	2023	CCTAGAGGGGTTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2082
Qy	2041	CCTTCTAGCGGTCTGCTCTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2100
Db	2083	CCTTCTAGCGGTCTGCTCTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2142
Qy	2101	GGTCCCATGTCATGAGTGTGCAACATGCAATGCTGTCTGTGTGTGTGTGTGTGTGTGTGT	2160
Db	2143	GGTCCCATGTCATGAGTGTGCAACATGCAATGCTGTCTGTGTGTGTGTGTGTGTGTGTGT	2200

QY 2161 AAAAAACACAGCCCTCTCTTTTCAGAGGAAAGGGGCCCTGAGGTGCCAGCTGTGCTCGGTT 2220  
Db AGAGACACAGCCCTCTCTTTTCAGAGGAAAGGGGCCCTGAGGTGCCAGCTGTGCTCGGTT 2260  
QY 2221 AGGGGTGGGGTCGGCCCTTCCAGGGCCAGAGGAGGAGGTTCCCTCTCTGTGTGCT 2280  
Db AGGGGTGGGGTCGGCCCTTCCAGGGCCAGAGGAGGAGGTTCCCTCTCTGTGTGCT 2320

QY 2281 GCTTGAAGTCTTAGAGGAAATAAAAAAGGGAAGTGAAGAAAAA 2326  
Db GCTTGAAGTCTTAGAGGAAATAAAAAAGGGAAGTGAAGAAAAA 2366

RESULT 3  
AB103033  
LOCUS AB103033 2525 bp mRNA linear PRI 29-OCT-2003  
DEFINITION Homo sapiens lat3 mRNA for L-type amino acid transporter 3,  
complete cds.  
ACCESSION AB103033  
VERSION AB103033.1 GI:38015937  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1  
Babu, E., Kanai, Y., Chairoungdua, A., Kim, D. K., Iribe, Y.,  
Tangtroungsup, S., Jutabha, P., Li, Y., Ahmed, N., Sakamoto, S.,  
Anzai, N., Naganori, S. and Endou, H.  
Identification of a novel system L amino acid transporter  
structurally distinct from heterodimeric amino acid transporters  
J. Biol. Chem. 278 (44), 43838-43845 (2003)

2 (bases 1 to 2525)  
12930836  
Kanai, Y.  
Direct Submission  
Submitted (07-PEB-2003) Yoshikatsu Kanai, Kyorin University School  
of Medicine, Department of Pharmacology and Toxicology; 6-20-2  
Shinkawa, Mitaka, Tokyo 181-8611, Japan  
(E-mail: ykanai@kyorin-u.ac.jp, Tel: 81-422-47-5511 (ex. 3453),  
Fax: 81-422-79-1321)

FEATURES  
source  
1. .2525  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="FLC4"  
/cell\_type="hepatoma"  
/tissue\_type="liver"  
/clone\_lib="psPORT1 FLC4 human hepatoma cell xDNA"  
/country="Japan"  
1. .2525  
/genes="lat3"  
257. .1936  
/genes="lat3"  
/function="system L amino acid transporter"  
/codon\_start=1  
/product="L-type amino acid transporter 3"  
/protein\_id="BAD00152.1"  
/db\_xref="GI:38015938"

gene  
CDS  
1. .2525  
/translation="MAPTLQAVRRRWMACTAVLENLFFSAVLWGSLIILKNEG  
FYSSCTPASSNTTODEQRWPGCDQDEMLNIGFTIGSFVLSATLPLGLIMDRFG  
PPRVLVGSACTASCTLMALASRDVEALSPILFALSUNGFGGCLIFTSILTNMF  
GNLRSTLMALMIGSYASSAITPGIKLIYDAGVAFVIMFTWSGLACLIFLNLNMF  
IEAFPAPEVNTTKLISGLADHKVTDGLFYHTVMTGRLSKQSPLEDSGDAFM  
SPQDVRGTSENPERSVPLRKSCTPFLWSLLTMGMTQLRIIFYMAAVNKMLEYLV  
GQOHEVNEQOQVAVTGVFSYVFGAMLLCLTCLPLIGIMDWRIKDCVDAPQTGT  
VLGDARDGVATSIIRPYCKIOKLTAISAFITLWLLVGRGITCTLNNHLIQVTFV  
LHTVIRGFHSACGSLIYAAPVFSRHFRTGLTGLQSLISAVFALLQQLPFLMNVGPKLGE  
PFWNLGLLFLSLLGLPLPSLYLFFYRRLQQEYAAANGMGLKVLGSGSEVTA"

ORIGIN

Query Match 96.1%; Score 2234.6; DB 9; Length 2525;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2252; Conservative 0; Mismatches 9; Indels 2; Gaps 1;  
QY 64 GCTCGGGCCACGATGGCCCCCACCCTGCAACAGGGGTACCGAGGCGTGGTGGATGCG 123  
Db GCTCGGGCCACGATGGCCCCCACCCTGCAACAGGGGTACCGAGGCGTGGTGGATGCG 303  
QY 124 CTGCAACGGCTGTGCTGGAGAACCTCTTCTCTGTCTGTACTCTCTGGGCTGGGGCTCCCT 183  
Db CTGCAACGGCTGTGCTGGAGAACCTCTTCTCTGTCTGTACTCTCTGGGCTGGGGCTCCCT 363  
QY 184 GTTGATCATCTTGAAGAACGAGGGCTTTATTTCCAGCACGTGCCCCAGCTGAGAGCAGCAC 243  
Db GTTGATCATCTTGAAGAACGAGGGCTTTATTTCCAGCACGTGCCCCAGCTGAGAGCAGCAC 423  
QY 244 CAACACACCCAGGATGAGCAGGCGAGGTGGCCAGGCTGTGACCAGCAGCAGAGATGCT 303  
Db CAACACACCCAGGATGAGCAGGCGAGGTGGCCAGGCTGTGACCAGCAGCAGAGATGCT 483  
QY 304 CAACCTGGGCTTCAACATTGGTTCTTCTGTCTGCTCAGCGCCACACCCTGCCACTGGGGAT 363  
Db CAACCTGGGCTTCAACATTGGTTCTTCTGTCTGCTCAGCGCCACACCCTGCCACTGGGGAT 543  
QY 364 CCTCATGGACCGCTTTGGCCCCCGACCCCGTGGCTGGTGGCAGTGCCTGTCTCACTGC 423  
Db CCTCATGGACCGCTTTGGCCCCCGACCCCGTGGCTGGTGGCAGTGCCTGTCTCACTGC 603  
QY 424 GTCTCGACCTCATGCGCCCTGGCTCCGGGACGTGGAGAGTCTGTCTCGGTGATATT 483  
Db GTCTCGACCTCATGCGCCCTGGCTCCGGGACGTGGAGAGTCTGTCTCGGTGATATT 663  
QY 484 CTGCGCGTGTCCCTGAATGGCTTTGGTGGCATCTGCCTAAAGTTCACCTTCACTCAGCT 543  
Db CTGCGCGTGTCCCTGAATGGCTTTGGTGGCATCTGCCTAAAGTTCACCTTCACTCAGCT 723  
QY 544 GCCCAACATGTTGGGAACCTCGCTCCACGTTAAATGGCCCTCATGATGGCTTCTTACGC 603  
Db GCCCAACATGTTGGGAACCTCGCTCCACGTTAAATGGCCCTCATGATGGCTTCTTACGC 783  
QY 604 CTCCTTCTGCCATTAGTTCCTCCAGGAATCAAGCTGATCTACGATGCCGGTGGCCTTCGT 663  
Db CTCCTTCTGCCATTAGTTCCTCCAGGAATCAAGCTGATCTACGATGCCGGTGGCCTTCGT 843  
QY 664 GGTTCATCATGTTCACTGGTCTGGCTGGCCCTGCCTTATCTTCTTGAACATGCACCTCAA 723  
Db GGTTCATCATGTTCACTGGTCTGGCTGGCCCTGCCTTATCTTCTTGAACATGCACCTCAA 903  
QY 724 CTGGCCCATCGAAGCCTTTCTGCCCCCTGAGGAAGTCAATTACAGGAAGAAGATCAAGCT 783  
Db CTGGCCCATCGAAGCCTTTCTGCCCCCTGAGGAAGTCAATTACAGGAAGAAGATCAAGCT 963  
QY 784 GAGTGGGCTGGCCCTGGACCAAGGTGACAGTGACCTCTTCTTACACCATGTGACCAC 843  
Db GAGTGGGCTGGCCCTGGACCAAGGTGACAGTGACCTCTTCTTACACCATGTGACCAC 1023  
QY 844 CATGGGCCAGAGGCTCAGCCAGAGGCCCCCGAGCTGGAGGACGGTTCGGATGCCCTTCAT 903  
Db CATGGGCCAGAGGCTCAGCCAGAGGCCCCCGAGCTGGAGGACGGTTCGGATGCCCTTCAT 1083  
QY 904 GTCACCCCAAGGATGTTTCGGGGCACCCTCAGAAAAACCTTCTCAGAGAGTCTGTCCCCCTTACG 963  
Db GTCACCCCAAGGATGTTTCGGGGCACCCTCAGAAAAACCTTCTCAGAGAGTCTGTCCCCCTTACG 1143  
QY 964 CAAGAGCCTCTGCTCCCCACATTTCTGTGGAGCCTCTCTCAACATGGGGCATGACCAGCT 1023  
Db CAAGAGCCTCTGCTCCCCACATTTCTGTGGAGCCTCTCTCAACATGGGGCATGACCAGCT 1203  
QY 1024 GCGGATCATCTTCTACATGGCTGTGTGAACAGAGTCTGGAGTACCTTGTGACTGTGTGG 1083  
Db GCGGATCATCTTCTACATGGCTGTGTGAACAGAGTCTGGAGTACCTTGTGACTGTGTGG 1263  
QY 1084 CCAGGAGCATGAGACAAATGAACAGCAACAAAGGGTGGCAGAGACAGTTGGTCTTCTACTC 1143

1264 CCAGAGCATGAGCAAAATGACACGACCAAAAGGTGGCAGAGACAGTTGGTCTACTC 1323  
 1144 CTCCTGTTTGGGGGCAATGACAGTGTGTGCTTCTCAGCTGCCCCCTCATTTGGGTACAT 1203  
 1324 CTCCTGTTTGGGGGCAATGACAGTGTGTGCTTCTCAGCTGCCCCCTCATTTGGGTACAT 1383  
 1204 CATGACATGGCGGATCAGGAGCTGGTGGAGCCGCCCACTCAGGACAGTGTCTCGGAGA 1263  
 1384 CATGACATGGCGGATCAGGAGCTGGTGGAGCCGCCCACTCAGGACAGTGTCTCGGAGA 1443  
 1264 TGCCAGGAGCGGGTGTCTACCAATTCATCAGACAGCTACTGCAAGATCCAAAGCT 1323  
 1444 TGCCAGGAGCGGGTGTCTACCAATTCATCAGACAGCTACTGCAAGATCCAAAGCT 1503  
 1324 CACCAATGACATCAGTGTCTCAGCTGACCAACCTGCTGTGGGTCTGGGTATGCTAC 1383  
 1504 CACCAATGACATCAGTGTCTCAGCTGACCAACCTGCTGTGGGTCTGGGTATGCTAC 1563  
 1384 CTGTCTCATCAACAACTTACACTCCAGTTTGTGACCTTTGCTGTCACACCATTTGTCG 1443  
 1564 CTGTCTCATCAACAACTTACACTCCAGTTTGTGACCTTTGCTGTCACACCATTTGTCG 1623  
 1444 AGGTTCCTTCCACTCAGCTGTGGAGTCTCTATGCTGCAGTGTCTCCATCCAACTT 1503  
 1624 AGGTTCCTTCCACTCAGCTGTGGAGTCTCTATGCTGCAGTGTCTCCATCCAACTT 1683  
 1504 TGGGACGCTGACAGGCTGAGTCTCTATGCTGCAGTGTCTCCATCCAACTT 1563  
 1684 TGGGACGCTGACAGGCTGAGTCTCTATGCTGCAGTGTCTCCATCCAACTT 1743  
 1564 ACTTTTCATGGGATGCTGGGACCTCTGAAAGGAGGAGGCTTCTGGGTGAATCTGGGCT 1623  
 1744 ACTTTTCATGGGATGCTGGGACCTCTGAAAGGAGGAGGCTTCTGGGTGAATCTGGGCT 1803  
 1624 CTGTCTATTTCTCAGTCTGGGATGCTGTGCTTCTTACCTTCTTATTTACCGTCCCG 1683  
 1804 CTGTCTATTTCTCAGTCTGGGATGCTGTGCTTCTTACCTTCTTATTTACCGTCCCG 1863  
 1684 GCTCAGAGGAGTACGCGCCCAATGGATGGGCCCACTGAAGTGTCTTAGCGGCTCTGA 1743  
 1864 GCTCAGAGGAGTACGCGCCCAATGGATGGGCCCACTGAAGTGTCTTAGCGGCTCTGA 1923  
 1744 GGTGACCGGATGAGTCTCAGACCAAGGAGCTGATGACAGGCAATCAAGGCTGAGC 1803  
 1924 GGTGACCGGATGAGTCTCAGACCAAGGAGCTGATGACAGGCAATCAAGGCTGAGC 1983  
 1804 AACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTGTAACATGCAATAGAGCCATGGCC 1863  
 1984 AACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTGTAACATGCAATAGAGCCATGGCC 2043  
 1864 GTAGATTTTAAATACCAAGAGAGTCTATTTTGTAAAGTGTGCAAAAGGAGGAAA 1923  
 2044 GTAGATTTTAAATACCAAGAGAGTCTATTTTGTAAAGTGTGCAAAAGGAGGAAA 2103  
 1924 AAAACCTTCAAAAAGGAGTGGCTTAAAGTCAAGCTTCAAGTGAAGACAGTCCCTATCCT 1983  
 2104 AAAACCTTCAAAAAGGAGTGGCTTAAAGTCAAGCTTCAAGTGAAGACAGTCCCTATCCT 2163  
 1984 AGAGGGGTTGAGCTTTCTTCTCTTGGGTGGAGGAGCCAGGCTGCTTATCTCTCT 2043  
 2164 AGAGGGGTTGAGCTTTCTTCTCTTGGGTGGAGGAGCCAGGCTGCTTATCTCTCT 2223  
 2044 TCTAGCGGTCTGCT 2103  
 2224 TCTAGCGGTCTGCT 2283  
 2104 CCCATGTGCGATGAGTGTGACAACTGCAATGTGTCTGTGTATGTGTGAATGTGAGAAA 2163  
 2284 CCCATGTGCGATGAGTGTGACAACTGCAATGTGTCTGTGTATGTGTGAATGTGAGAGA 2341  
 2164 AACACAGCCCTCTCTTTCAGAGGAAAGGGGCTGAGGTGCCAGCTGTCTCTGGGTAGG 2223

2342 GACACAGCCCTCTTTTCAGAGGAAAGGGGCTTCCAGGTGCCAGCTGTGTCTGGGTAGG 2401  
 2224 GGTGGGGGTGGGCGCCCTTCCAGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2283  
 2402 GGTGGGGGTGGGCGCCCTTCCAGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2461  
 2284 TGCAAGTCTTAGAGGAAATAAAAAGGGAAGTGAAGAAAAA 2326  
 2462 TGCAAGTCTTAGAGGAAATAAAAAGGGAAGTGAAGAAAAA 2504

RESULT 4  
 BD063631  
 LOCUS  
 DEFINITION Human proteins having transmembrane domains and DNAs encoding these proteins.  
 ACCESSION BD063631  
 VERSION BD063631.1 GI:22609234  
 KEYWORDS JP 2001508407-A/46.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2373)  
 AUTHORS Kato,S., Sekine,S., Kimura,T. and Kobayashi,M.  
 TITLE Human proteins having transmembrane domains and DNAs encoding these proteins  
 JOURNAL Patent: JP 2001508407-A 46 26-JUN-2001;  
 COMMENT SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC  
 OS Homo sapiens (human)  
 PN JP 2001508407-A/46  
 PD 26-JUN-2001  
 PF 07-NOV-1997 JP 1998522374  
 PR 13-NOV-1996 JP 8/301429  
 PI SEISHI KATO, SHINGO SEKINE, TOMOKO KIMURA, MIDORI KOBAYASHI PC  
 C12N15/12,C07K14/705,C12N5/10,C12N15/57,C12N9/48,C12N9/14, PC  
 C12N15/55  
 CC Strandedness: Double;  
 CC Topology: Linear;  
 FH Key  
 FT CDS  
 Location/Qualifiers  
 1..2373  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

FEATURES  
 source  
 ORIGIN

Query Match 95.2%; Score 2214.6; DB 6; Length 2373;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 2243; Conservative 0; Mismatches 9; Indels 3; Gaps 2;

QY 64 GCTGGGGCCAGCATGGCCCCCAGCTGCAACAGCGGTACCGGAGGCGCTGGTGGATGCG 123  
 DB 121 GCTGGGGCCAGCATGGCCCCCAGCTGCAACAGCGGTACCGGAGGCGCTGGTGGATGCG 180  
 QY 124 CTGCACGCTGTGTGGAGAACCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 183  
 DB 181 CTGCACGCTGTGTGGAGAACCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
 QY 184 GTTGATCTATTGAAAGAACGAGGGGCTTCTATTTCCAGCACGTCGCCAGCTGAGAGAGCAC 243  
 DB 241 GTTGATCTATTGAAAGAACGAGGGGCTTCTATTTCCAGCACGTCGCCAGCTGAGAGAGCAC 300  
 QY 244 CAAACCCACCCAGGATGAGCAGCGCAGGTGGCAGGCTGTGACCCAGGAGGAGGATGCT 303  
 DB 301 CAAACCCACCCAGGATGAGCAGCGCAGGTGGCAGGCTGTGACCCAGGAGGAGGATGCT 360  
 QY 304 CAACTGGGCTTACCACTTGGTTCCTTCTGTCTCAGCGCCACCACTGCGGAT 363  
 DB 361 CAACTGGGCTTACCACTTGGTTCCTTCTGTCTCAGCGCCACCACTGCGGAT 420  
 QY 364 CCTCATGACCGCTTTGGCCCCCGACCCCGTGGGCTGGTTGGCAGTGCCTTCACTGC 423









1363 GCTTGTGGTGTGGCATCACCTGCTCATCAACAACTTACACCTCCAGTTTGTGACCTT 1422  
 Db GCTTGTGGTGTGGCATCACCTGCTCATCAACAACTTACACCTCCAGTTTGTGACCTT 1200  
 1423 TGTCTGCACACCAATGTTTCGAGGTTTCTTCCACTCAGCCTGTGGAGTCTCTATGTCG 1482  
 Db TGTCTGCACACCAATGTTTCGAGGTTTCTTCCACTCAGCCTGTGGAGTCTCTATGTCG 1260  
 1483 AGTGTTCATCCATCCACCACTTGTGGACGCTGACAGGCTGACGCTCCCTCATCAGTGTGT 1542  
 Db AGTGTTCATCCATCCACCACTTGTGGACGCTGACAGGCTGACGCTCCCTCATCAGTGTGT 1320  
 1543 GTTCCGCTTGTCTCAGCAGCACTTTTTCATGCGCATGTGGACCCCTGAAAGAGAGACC 1602  
 Db GTTCCGCTTGTCTCAGCAGCACTTTTTCATGCGCATGTGGACCCCTGAAAGAGAGACC 1380  
 1603 CTTCTGGGTGAATCTGGGCTCTCTCTATTTCTCACTCTCTGGGATTCCTGTTCCTTCTTA 1662  
 Db CTTCTGGGTGAATCTGGGCTCTCTCTATTTCTCACTCTCTGGGATTCCTGTTCCTTCTTA 1440  
 1663 CTTCTTTATTTACCTGCTGGGCTCTCAGCAGGAGTACGCGCCATGGATGGGCTCACT 1722  
 Db CTTCTTTATTTACCTGCTGGGCTCTCAGCAGGAGTACGCGCCATGGATGGGCTCACT 1500  
 1723 GAAGTGTCTTAGCGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGATG 1782  
 Db GAAGTGTCTTAGCGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGATG 1560  
 1783 ACAGGCAATCAAGGCTGAGCAACCAAAAGGAGTGTCCCATATGCTTTTCTACCTGTA 1842  
 Db ACAGGCAATCAAGGCTGAGCAACCAAAAGGAGTGTCCCATATGCTTTTCTACCTGTA 1620  
 1843 CATGACATAGAGCCTGAGTGTGATTTAATAATACAGAGAGTGTCTATTTTGTAA 1902  
 Db CATGACATAGAGCCTGAGTGTGATTTAATAATACAGAGAGTGTCTATTTTGTAA 1680  
 1903 AGACTGCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1962  
 Db AGACTGCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740  
 1963 ACTGAAGACAGTCCCTATCTTAGAGGGTGTGAGTCTTCTTCTCTGGGTTGGAGGAGA 2022  
 Db ACTGAAGACAGTCCCTATCTTAGAGGGTGTGAGTCTTCTTCTCTGGGTTGGAGGAGA 1800  
 2023 CCAGGCTGCTTATCTCTTCTAGCGGTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 2082  
 Db CCAGGCTGCTTATCTCTTCTAGCGGTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1860  
 2083 CAACACAGGCTACCCCTGAGGTCCCATGTGCCATGAGTGTGCAACATGCAATGTGTCTG 2142  
 Db CAACACAGGCTACCCCTGAGGTCCCATGTGCCATGAGTGTGCAACATGCAATGTGTCTG 1918  
 2143 TGTATGTGAATGTGAGAGAAACACAGCCCTCTCTTTCAGAGGAAAGGGGCTCAGGTG 2202  
 Db TGTATGTGAATGTGAGAGAGACACAGCCCTCTCTTTCAGAGGAAAGGGGCTCAGGTG 1978  
 2203 CCAGCTGTGCTCTGAGGTTAGGGGTTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGG 2262  
 Db CCAGCTGTGCTCTGAGGTTAGGGGTTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGG 2038  
 2263 TCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2317  
 Db TCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2039

RESULT 7  
 LOCUS BD158342 2091 bp DNA linear PAT 17-JAN-2003  
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION BD158342  
 VERSION BD158342.1 GI:27864100  
 KEYWORDS JP 2002191363-A/13185

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 2093)  
 AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
 TITLE Primer for synthesizing full-length cDNA and use thereof  
 JOURNAL Patent: JP 2002191363-A 13185 09-JUL-2002;  
 HELIX RESEARCH INSTITUTE  
 COMMENT OS Homo sapiens (human)  
 PN JP 2002191363-A/13185  
 PD 09-JUL-2002  
 PF 28-JUL-2000 JP 2000280990  
 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU  
 PI SAITO,  
 PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,  
 PI KEIICHI NAGAI,TETSUJI OTSUKI  
 PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10,PC  
 C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10,CC  
 PC C12P21/02,C12P1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00,CC  
 Primer for synthesizing full-length cDNA and use thereof FH Key  
 Location/Qualifiers  
 FT CDS (77)..(1531).  
 source 1..2093  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 88.8%; Score 2066.6; DB 6; Length 2093;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 2084; Conservative 0; Mismatches 9; Indels 2; Gaps 1;  
 QY 223 GTGCCAGCTGAGAGCAGCACCACACCCAGATGAGCAGCGCAGGTGCGCAGGCTG 282  
 Db 1 GTGCCAGCTGAGAGCAGCACCACACCCAGATGAGCAGCGCAGGTGCGCAGGCTG 60  
 QY 283 TGACAGCAGGACGAGATGCTCAACCTGGGCTTCAACATTTGGTCTTCGTCAGCGC 342  
 Db 61 TGACAGCAGGACGAGATGCTCAACCTGGGCTTCAACATTTGGTCTTCGTCAGCGC 120  
 QY 343 CACACCTGCTGCTGGGATCTCATGGACCGCTTTGGGCCCCGACCCGCTGGCTGT 402  
 Db 121 CACACCTGCTGCTGGGATCTCATGGACCGCTTTGGGCCCCGACCCGCTGGCTGT 180  
 QY 403 TGGCAGTGCCTGCTTCACTGCGTCTGACCCCTCATGSCCTTGGCCTCCCGGAGCTGGA 462  
 Db 181 TGGCAGTGCCTGCTTCACTGCGTCTGACCCCTCATGSCCTTGGCCTCCCGGAGCTGGA 240  
 QY 463 AGCTGTCTCCGTTGATTTCTTGGGCTGTCTCTGATGGCTTTGGTGGCATCTGCT 522  
 Db 241 AGCTGTCTCCGTTGATTTCTTGGGCTGTCTCTGATGGCTTTGGTGGCATCTGCT 300  
 QY 523 AACGTTCACTTCACTGCGTCCCAACATGTTTGGGAACCTGCGCTCCAGTTAATGCG 582  
 Db 301 AACGTTCACTTCACTGCGTCCCAACATGTTTGGGAACCTGCGCTCCAGTTAATGCG 360  
 QY 583 CCTCATGATGGCTCTTACGCTCTTCTGCAATACGTTCCAGGAATCAAGATGATCTA 642  
 Db 361 CCTCATGATGGCTCTTACGCTCTTCTGCAATACGTTCCAGGAATCAAGATGATCTA 420  
 QY 643 CGATGCCGTTGGCTTCTGTTGATCATGTTTCACTGTTCTGGCTGGCTGGCTTAT 702  
 Db 421 CGATGCCGTTGGCTTCTGTTGATCATGTTTCACTGTTCTGGCTGGCTGGCTTAT 480  
 QY 703 CTTTCTGAATGACCTCAACTGCGCATCGAAGCTTTCTGCTCCCTGAGGAAGTCAA 762  
 Db 481 CTTTCTGAATGACCTCAACTGCGCATCGAAGCTTTCTGCTCCCTGAGGAAGTCAA 540  
 QY 763 TTACACGAAGAAGATCAAGCTGAGTGGGCTGGGCTTGGACCAACAGGTGACGCT 822



541	Db	 TTACAGAAAGATCAAGCTGAGTGGCTGGCCCTGGACCAAGGTGACAGTGACCT	600
823	Qy	 CTTCTACACCCATGTGACACCACTGGCCAGAGGCTCAGCCAGAGGCCCCAGCCTGGA	882
601	Db	 CTTCTACACCCATGTGACCACTGGCCAGAGGCTCAGCCAGAGGCCCCAGCCTGGA	660
883	Qy	 GGACGGTTCCGATGCCCTTCACTGTCAACCCAGGATGTTGGGGCACCTCGAAGAACCTTCC	942
661	Db	 GGACGGTTCCGATGCCCTTCACTGTCAACCCAGGATGTTGGGGCACCTCGAAGAACCTTCC	720
943	Qy	 TGAGAGTCTGTCTCCCTTACGGAAGGCTCTGCTCTCCCACTTTCCTGTGAGAGCCTCCT	1002
721	Db	 TGAGAGTCTGTCTCCCTTACGGAAGGCTCTGCTCTCCCACTTTCCTGTGAGAGCCTCCT	780
1003	Qy	 CACCATGGGCATGACCCAGCTGCGGATCATCTTCTATCATGGCTGCTGTGAACAAGATGCT	1062
781	Db	 CACCATGGGCATGACCCAGCTGCGGATCATCTTCTATCATGGCTGCTGTGAACAAGATGCT	840
1063	Qy	 GGAGTACCTTGTGACTGCTGGTGGCCAGGAGCATGAGACAAATGAACAGAAAGGTGGC	1122
841	Db	 GGAGTACCTTGTGACTGCTGGTGGCCAGGAGCATGAGACAAATGAACAGAAAGGTGGC	900
1123	Qy	 AGAGACAGTTGGGTTCTACTCTCCGTTCTGGGGGCCATGACAGCTGTGTGCTTCTCAC	1182
901	Db	 AGAGACAGTTGGGTTCTACTCTCCGTTCTGGGGGCCATGACAGCTGTGTGCTTCTCAC	960
1183	Qy	 CTGCCCCCTCATTTGGCTTACATGAGCTGGCGGATCAAGGACTGGGTGGAGCCGCCAAC	1242
961	Db	 CTGCCCCCTCATTTGGCTTACATGAGCTGGCGGATCAAGGACTGGGTGGAGCCGCCAAC	1020
1243	Qy	 TCAGGGCACTGTCTCTCGAGATGCCAGGACGGGTTGCTTACCAATTCATCAGACCAAG	1302
1021	Db	 TCAGGGCACTGTCTCTCGAGATGCCAGGACGGGTTGCTTACCAATTCATCAGACCAAG	1080
1303	Qy	 CTACTGCAAGATCCAAAGCTCAACCAATGCCATCAGTGCCTTCAACCTGACCAACCTGCT	1362
1081	Db	 CTACTGCAAGATCCAAAGCTCAACCAATGCCATCAGTGCCTTCAACCTGACCAACCTGCT	1140
1363	Qy	 GCTTGTGGGTTTGGCATCACTGTCTCATCAACACTTACACTCCAGTTTCTGACCTT	1422
1141	Db	 GCTTGTGGGTTTGGCATCACTGTCTCATCAACACTTACACTCCAGTTTCTGACCTT	1200
1423	Qy	 TGTCTCTGCACACATTTCTCGAGTTTCTTCCACTCAGCCTGTGGGAGTCTCTATGCTGC	1482
1201	Db	 TGTCTCTGCACACATTTCTCGAGTTTCTTCCACTCAGCCTGTGGGAGTCTCTATGCTGC	1260
1483	Qy	 AGTGTTCCTCATCAAACACTTTTGGGACGCTGACAGGCTCGAGTGCCTCATAGTGCCTGT	1542
1261	Db	 AGTGTTCCTCATCAAACACTTTTGGGACGCTGACAGGCTCGAGTGCCTCATAGTGCCTGT	1320
1543	Qy	 GTTTGGCCTTCTTCAGCAGCACTTTTTCATGCGGATGGTGGGACCCCTGGAAGGAGAGCC	1602
1321	Db	 GTTTGGCCTTCTTCAGCAGCACTTTTTCATGCGGATGGTGGGACCCCTGGAAGGAGAGCC	1380
1603	Qy	 CTTCTGGGTGAATCTGGGCCCTCTGCTATTCTCACTCCTGGGATCTCTGTGCTTCTCCTA	1662
1381	Db	 CTTCTGGGTGAATCTGGGCCCTCTGCTATTCTCACTCCTGGGATCTCTGTGCTTCTCCTA	1440
1663	Qy	 CCTCTTCTATTACCGTGCCCGGCTCCAGCAGGAGTAGCGCGCAATGGGATGGGCCCACT	1722
1441	Db	 CCTCTTCTATTACCGTGCCCGGCTCCAGCAGGAGTAGCGCGCAATGGGATGGGCCCACT	1500
1723	Qy	 GAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCCAGGGACCTGGATG	1782
1501	Db	 GAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCCAGGGACCTGGATG	1560
1783	Qy	 ACAGGCAATCAAGSCCTGAGCAACCAAAAGAGTGGCCCACTATGCTTTTCTACCTGTAA	1842
1561	Db	 ACAGGCAATCAAGSCCTGAGCAACCAAAAGAGTGGCCCACTATGCTTTTCTACCTGTAA	1620
1843	Qy	 CATGCAATAGAGCCATGGCCGTAGATTTTATAAATACCAAGAGAAGTTCTATTTTGTAA	1902

Db	1621	CATGCAATAGAGCCATGCGCGTAGATTTATTAATAACCAAGAGAGTCTCTATTTTTGTGTA	1688
QY	1903	AGACTCAAAAAGGAGGAAAAAAACCTTTCAAAAACGCCCCCTTAAGCTCAACGCTCCCATTCG	1962
Db	1681	AGACTGCAAAAAGGAGGAAAAAAACCTTTCAAAAACGCCCCCTTAAGCTCAACGCTCCCATTCG	1740
QY	1963	ACTGAAGACAGTCCCTTATCCTTAGAGGGGTGAGCTTTCTTCTCTTGGGTTCGAGGAGA	2022
Db	1741	ACTGAAGACAGTCCCTTATCCTTAGAGGGGTGAGCTTTCTTCTCTTGGGTTCGAGGAGA	1800
QY	2023	CCAGGGTGCCTCTTATCTCTTCTTAGGGGTCTGCCTCTGGTACCTCTTTGGGGGGATCGG	2082
Db	1801	CCAGGGTGCCTCTTATCTCTTCTTAGGGGTCTGCCTCTGGTACCTCTTTGGGGGGATCGG	1860
QY	2083	CAAAACAGGCTACCCCTGAGGTCCCATGTGCGATGAGTGTGCACAAATGCAATGTGTCTG	2142
Db	1861	CAAAACAGGCTACCCCTGAGGTCCCATGTGCGATGAGTGTGCACAAATGCAATGTGTCTG	1918
QY	2143	TGTATGTGTAATGTGAGAAAAACACAGCCCTCTTTTCAGAGAAAGGGCCCTGAGGTG	2202
Db	1919	TGTATGTGTAATGTGAGAGAGACACAGCCCTCTTTTCAGAGAAAGGGCCCTGAGGTG	1978
QY	2203	CCAGCTGTCTCTGGGTTAGGGGTTCGGGGTTCGGCCCTTCCAGAGGCCAGGAGGCGAGGT	2262
Db	1979	CCAGCTGTCTCTGGGTTAGGGGTTCGGGGTTCGGCCCTTCCAGAGGCCAGGAGGCGAGGT	2038
QY	2263	TCCTCTCTGGTCTCTCTGCTTGTCAAGTCTTAGAGGAAATAAAAAGGAAAGTGTGAG	2317
Db	2039	TCCTCTCTGGTCTCTCTGCTTGTCAAGTCTTAGAGGAAATAAAAAGGAAAGTGTGAG	2093
RESULT 8			
AK027732			
LOCUS			
DEFINITION	AK027732 2093 bp mRNA linear PRI 01-AUG-2002		
ACCESSION	Homo sapiens cDNA FLJ14826 fis, clone OVARC1000850, highly similar		
VERSION	to Homo sapiens PB39 mRNA.		
KEYWORDS	AK027732.1 GI:14042628		
SOURCE	oligo capping; fis (full insert sequence).		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 Isogai, T., Ota, T., Hayaashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagaatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.		
JOURNAL	NEDO human cDNA sequencing project		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 2093)		
TITLE	Isogai, T. and Otsuki, T.		
JOURNAL	Direct Submission		
COMMENT	Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.		
FEATURES	Location/Qualifiers		
source	1..2093		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="OVARC1000850"		
	/tissue type="ovary, tumor tissue"		

/clone_lib="OVARCI1" /note="cloning vector: pME18SFL3"									
ORIGIN									
Query Match 88.8%; Score 2066.6; DB 9; Length 2093; Best Local Similarity 99.5%; Pred. No. 0; Matches 2084; Conservative 0; Mismatches 9; Indels 2; Gaps 1;									
QY	223	GTGCCAGCTGAGACGACGACCAACACACCCAGGATGAGACGCGCAGGTGGCCAGGCTG	282						
DB	1	GTGCCAGCTGAGACGACGACCAACACACCCAGGATGAGACGCGCAGGTGGCCAGGCTG	60						
QY	283	TGACACGAGCAGAGAGTCTCAACTGGGCTTACCAATGGTTGCTTCCTTCGCTCAGCGC	342						
DB	61	TGACACGAGCAGAGAGTCTCAACTGGGCTTACCAATGGTTGCTTCCTTCGCTCAGCGC	120						
QY	343	CACCAACCTGCCACTGGGGATCCTCATGGAACGCTTTTGGCCCCCGACCCGCTGGCTGGT	402						
DB	121	CACCAACCTGCCACTGGGGATCCTCATGGAACGCTTTTGGCCCCCGACCCGCTGGCTGGT	180						
QY	403	TGGCAGTGCCTGCTTCACTGCGTCTGCACCTCATGGCCCTGGCCTCCCGGAGCTGA	462						
DB	181	TGGCAGTGCCTGCTTCACTGCGTCTGCACCTCATGGCCCTGGCCTCCCGGAGCTGA	240						
QY	463	AGCTCTGTCTCCGTTGATATTCTTGGCGCTGTCCCTGAATGGCTTTGGTGCACTGCT	522						
DB	241	AGCTCTGTCTCCGTTGATATTCTTGGCGCTGTCCCTGAATGGCTTTGGTGCACTGCT	300						
QY	523	AACGTTCACTTCACTCAGCTGCCCAACATGTTTGGGAACCTGCGCTCAAGTTAATGGC	582						
DB	301	AACGTTCACTTCACTCAGCTGCCCAACATGTTTGGGAACCTGCGCTCAAGTTAATGGC	360						
QY	583	CCTCATGATTGGCTCTTACGCTCTTCTGCCATTAGTTCCTCAGGAATCAAGCTCATCTA	642						
DB	361	CCTCATGATTGGCTCTTACGCTCTTCTGCCATTAGTTCCTCAGGAATCAAGCTCATCTA	420						
QY	643	CGATCCGGTGTGGCTTTCGTGGTCACTCATGTTTACCTGGTCTGGCTGGCTGCTTAT	702						
DB	421	CGATCCGGTGTGGCTTTCGTGGTCACTCATGTTTACCTGGTCTGGCTGGCTGCTTAT	480						
QY	703	CTTTCTGAATCGACCTCAACTGGCCCATCGAAGCTTTCCTGCCCTGAGGAAGTCAA	762						
DB	481	CTTTCTGAATCGACCTCAACTGGCCCATCGAAGCTTTCCTGCCCTGAGGAAGTCAA	540						
QY	763	TTACACGAAGAAGATCAAGCTGAGTGGGCTGGCCCTGGACCAACAAGGTGACAGTCACT	822						
DB	541	TTACACGAAGAAGATCAAGCTGAGTGGGCTGGCCCTGGACCAACAAGGTGACAGTCACT	600						
QY	823	CTTTACACCCATGTGACCAACCATGGCCAGAGGCTCAGCCAGAGAGGCCCCAGCCTGA	882						
DB	601	CTTTACACCCATGTGACCAACCATGGCCAGAGGCTCAGCCAGAGAGGCCCCAGCCTGA	660						
QY	883	GGACGGTTGGATGCTTCACTGACCCAGGATGTTTGGGGCACTCAGAAACCTTCC	942						
DB	661	GGACGGTTGGATGCTTCACTGACCCAGGATGTTTGGGGCACTCAGAAACCTTCC	720						
QY	943	TCGAGGTTCTGCTCCCTTACGCAAGAGCTCTGCTCCCACTTCTCTGTGGAGCTCT	1002						
DB	721	TGAGAGGTTCTGCTCCCTTACGCAAGAGCTCTGCTCCCACTTCTCTGTGGAGCTCT	780						
QY	1003	CACCATGGGCATGACCCAGCTCGGATCATCTTCTACATGGCTGTGTGAAACAAGATGCT	1062						
DB	781	CACCATGGGCATGACCCAGCTCGGATCATCTTCTACATGGCTGTGTGAAACAAGATGCT	840						
QY	1063	GGAGTACCTTGTGATGTGGTGGCCAGGATGAGACAAATGAAACAGCAACAAAGGTGGC	1122						
DB	841	GGAGTACCTTGTGATGTGGTGGCCAGGATGAGACAAATGAAACAGCAACAAAGGTGGC	900						
QY	1123	AGAGACAGTGTGGTCTTACTCTCTCGGTCTTTCGGGGCCATGACAGTGTGTGCTTCTCAC	1182						
DB	901	AGAGACAGTGTGGTCTTACTCTCTCGGTCTTTCGGGGCCATGACAGTGTGTGCTTCTCAC	960						
QY	1183	CTGCCCTCTCATTTGGCTACATCATGACTGGCGGATCAAGGACTGCGTGGACGCCCAAC	1242						

DB	961	CTGCCCTCTCATTTGGCTACATCATGAGCTGCGGATCAAGGACTGCGTGGACGCCCAAC	1020						
QY	1243	TCAGGGCACTGTCTCGGAGATGCCAGGACGGGTTGCTACCAAAATCCATCAGACCAACG	1302						
DB	1021	TCAGGGCACTGTCTCGGAGATGCCAGGACGGGTTGCTACCAAAATCCATCAGACCAACG	1080						
QY	1303	CTACTGCAAGATCCAAAGCTCACCAATGGCATCAGTGCCTTCCACCTGACCAACCTGCT	1362						
DB	1081	CTACTGCAAGATCCAAAGCTCACCAATGGCATCAGTGCCTTCCACCTGACCAACCTGCT	1140						
QY	1363	GCTTGTGGGTTTGGCATCACCTCTCTCAACAACTTACCTCAGTTTGTGACCTT	1422						
DB	1141	GCTTGTGGGTTTGGCATCACCTCTCTCAACAACTTACCTCAGTTTGTGACCTT	1200						
QY	1423	TGTCCTGCACACCAATGTTGAGGTTTCTTCCACTCAGCCTGCGGAGTCTCTATGCTGC	1482						
DB	1201	TGTCCTGCACACCAATGTTGAGGTTTCTTCCACTCAGCCTGCGGAGTCTCTATGCTGC	1260						
QY	1483	AGTGTTCCTCATCCAAACCACTTTGGGACGCTGACAGGCTGCGAGTCCCTCATCAGTGTGT	1542						
DB	1261	AGTGTTCCTCATCCAAACCACTTTGGGACGCTGACAGGCTGCGAGTCCCTCATCAGTGTGT	1320						
QY	1543	GTTGCGCTGTCTCAGCAGGCACTTTTTCATGGCGATGGTGGGACCCCTGAAAGGAGAGCC	1602						
DB	1321	GTTGCGCTGTCTCAGCAGGCACTTTTTCATGGCGATGGTGGGACCCCTGAAAGGAGAGCC	1380						
QY	1603	CTTCTGGGTGAATCTGGGCTCTGCTATTCTCACTCCTGGGATTCCTGTTGCCCTTCTCTA	1662						
DB	1381	CTTCTGGGTGAATCTGGGCTCTGCTATTCTCACTCCTGGGATTCCTGTTGCCCTTCTCTA	1440						
QY	1663	CCTCTTCTATTATCCGCTGCCGCTCCAGCAGGAGTAGCCGCCCAATGGGATGGGCCACT	1722						
DB	1441	CCTCTTCTATTATCCGCTGCCGCTCCAGCAGGAGTAGCCGCCCAATGGGATGGGCCACT	1500						
QY	1723	GAAAGTGTCTAGCGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGAGCCTGGATG	1782						
DB	1501	GAAAGTGTCTAGCGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGAGCCTGGATG	1560						
QY	1783	ACAGGCAATCAAGGCTTGAGCAACCAAAAGAGTGGCCCATATGCTTTCTACTGTAA	1842						
DB	1561	ACAGGCAATCAAGGCTTGAGCAACCAAAAGAGTGGCCCATATGCTTTTCTACTGTAA	1620						
QY	1843	CATGCACATAGAGCATGGCCGTAGATTATAAATACCAAGAGAAAGTTCTATTTTGTAA	1902						
DB	1621	CATGCACATAGAGCATGGCCGTAGATTATAAATACCAAGAGAAAGTTCTATTTTGTAA	1680						
QY	1903	AGACTGCAAAAAGAGGAGGAAAAAACCCTTCAAAAACGCCCCCTTAAGTCAACGCTCCATG	1962						
DB	1681	AGACTGCAAAAAGAGGAGGAAAAAACCCTTCAAAAACGCCCCCTTAAGTCAACGCTCCATG	1740						
QY	1963	ACTGAAGACAGTCCCTATCCTAGAGGGGTTGAGCTTTCTCTCTCTGGTTGGAGGAGA	2022						
DB	1741	ACTGAAGACAGTCCCTATCCTAGAGGGGTTGAGCTTTCTCTCTCTGGTTGGAGGAGA	1800						
QY	2023	CCAGGGTCCCTCTTATCTCTTCTAGCGGTCTGCTCCTCTGCTACCTCTTGGGGGATCGG	2082						
DB	1801	CCAGGGTCCCTCTTATCTCTTCTAGCGGTCTGCTCCTCTGCTACCTCTTGGGGGATCGG	1860						
QY	2083	CAAAACAGGCTACCCCTGAGGTCCCATGTCATGATGTGCAACAATGCTATGTGCTG	2142						
DB	1861	CAAAACAGGCTACCCCTGAGGTCCCATGTCATGATGTGCAACAATGCTATGTGCTG	1918						
QY	2143	TGATGTGTGAATGTGAGAAAAACACAGCCCTCTTTTCAGAAAGGAAAGGGGCTGAGGTG	2202						
DB	1919	TGATGTGTGAATGTGAGAGAGACACAGCCCTCTTTTCAGAAAGGAAAGGGGCTGAGGTG	1978						
QY	2203	CCAGCTGTCTCTGGGTTAGGGGTTGGGGTTCGGGCTTCGGCCCTTCCAGGGCCAGGAAGCAGT	2262						
DB	1979	CCAGCTGTCTCTGGGTTAGGGGTTGGGGTTCGGGCTTCGGCCCTTCCAGGGCCAGGAAGCAGT	2038						
QY	2263	TCCCTCTCTGTGCTGCTGTGAGTCTTAGAGGAATATAAAGGGAAGTGTAG 2317							

Db 2039 TCCTCTCTGGTCTGCTGCTTCAAGTCTTAGAGGAATAAAAGGAAGTGAG 2093

RESULT 9  
LOCUS BD063606 1677 bp DNA linear PAT 27-AUG-2002  
DEFINITION Human proteins having transmembrane domains and DNAs encoding these proteins.  
ACCESSION BD063606  
VERSION BD063606.1 GI:22609209  
KEYWORDS JP 2001508407-A/21.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1677)  
AUTHORS Kato,S., Sekine,S., Kimura,T. and Kobayashi,M.  
TITLE Human proteins having transmembrane domains and DNAs encoding these proteins  
JOURNAL Patent: JP 2001508407-A 21 26-JUN-2001;  
COMMENT SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC  
OS Homo sapiens (human)  
PN JP 2001508407-A/21  
PD 26-JUN-2001  
PF 07-NOV-1997 JP 1998522374  
PR 13-NOV-1996 JP 8/301429  
PI SEISHI KATO, SHINGO SEKINE, TOMOKO KIMURA, MIDORI KOBAYASHI PC  
C12N15/12, C07K14/705, C12N5/10, C12N15/57, C12N9/48, C12N9/14, PC  
C12N15/55  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
1. .1677  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 72.1%; Score 1677; DB 6; Length 1677;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 ATGGCCCCACAGCTGCAACAGGGGTACCGAGCGCTGGTGGATGCGCTCGACGGCTGTG 136  
DB 1 ATGGCCCCACAGCTGCAACAGGGGTACCGAGCGCTGGTGGATGCGCTCGACGGCTGTG 60

QY 137 CTGGAGAACTCTTCTCTCTGCTGTACTCTCTGGGCTGGGGCTCCCTGTGTGATCTCTG 196  
DB 61 CTGGAGAACTCTTCTCTCTGCTGTACTCTCTGGGCTGGGGCTCCCTGTGTGATCTCTG 120

QY 197 AAGAACGAGGGCTTCTATTCAGCAGCTGCGCCAGCTGAGAGCAGCAACACACCCAG 256  
DB 121 AAGAACGAGGGCTTCTATTCAGCAGCTGCGCCAGCTGAGAGCAGCAACACACCCAG 180

QY 257 GATGACGAGCGAGTGGCGAGCTGTGACACGAGCAGAGATGCTCAACCTGGGCTTC 316  
DB 181 GATGACGAGCGAGTGGCGAGCTGTGACACGAGCAGAGATGCTCAACCTGGGCTTC 240

QY 317 ACCATTGGTTCCTTCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376  
DB 241 ACCATTGGTTCCTTCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

QY 377 TTGGCCCCGAGCCCTGGGCTGGTGGCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436  
DB 301 TTGGCCCCGAGCCCTGGGCTGGTGGCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

QY 437 ATGGCCCTGGCTCCCGGAGCGTGAAGTCTGCTCCGTTGATATTCCTGGGCTGTGC 496  
DB 361 ATGGCCCTGGCTCCCGGAGCGTGAAGTCTGCTCCGTTGATATTCCTGGGCTGTGC 420

QY 497 CTGAATGGCTTTGGTGGCATCTGCGCTAACGTTCACTTCACTCACTGCGTGCACCATGTTT 556

Db 421 CTGAATGGCTTTGGTGGCATCTGCGCTAACGTTCACTTCACTCACTGCGTGCACCATGTTT 480

QY 557 GGGAACTGCGCTCCAGCGTTAATGGCCCTCATGATGGCTCTTACGCGCTCTTCTGCGCATT 616

Db 481 GGGAACTGCGCTCCAGCGTTAATGGCCCTCATGATGGCTCTTACGCGCTCTTCTGCGCATT 540

QY 617 AGCTTCCCAAGGAATCAAGCTGATACGATGCGGCTGCGGCTTCTGCTGCTCATCATGTTTC 676

Db 541 AGCTTCCCAAGGAATCAAGCTGATACGATGCGGCTGCGGCTTCTGCTGCTCATCATGTTTC 600

QY 677 ACTGGTCTGGCTGCGCTGCGCTTATCTTTCTGAACTGCAACCTCACTCACTGCGGCTCATGAA 736

Db 601 ACTGGTCTGGCTGCGCTGCGCTTATCTTTCTGAACTGCAACCTCACTCACTGCGGCTCATGAA 660

QY 737 GCCTTTCCTGCGGCTGAGGAAGTCAATACAGGAAGATCAAGCTGAGTGGCTGGCC 796

Db 661 GCCTTTCCTGCGGCTGAGGAAGTCAATACAGGAAGATCAAGCTGAGTGGCTGGCC 720

QY 797 CTGGACCAAGGTCAGGTCAGCTTCTTACACCCATGTGACCACTGACCACTGCGGCTCATCTTC 856

Db 721 CTGGACCAAGGTCAGGTCAGCTTCTTACACCCATGTGACCACTGCGGCTCATCTTC 780

QY 857 CTCAGCAGAACGCGCCAGCTGAGGACGCTTGGATGCGCTTCACTGACACCCAGGAT 916

Db 781 CTCAGCAGAACGCGCCAGCTGAGGACGCTTGGATGCGCTTCACTGACACCCAGGAT 840

QY 917 GTTGGGGGACCTCAGAAACCTTCTGAGAGGCTGTCCTTACGAGAGGCTTACGAGAGGCTCTGC 976

Db 841 GTTGGGGGACCTCAGAAACCTTCTGAGAGGCTGTCCTTACGAGAGGCTCTGC 900

QY 977 TCCCCACCTTCTGTTGGAGCTCTCTACCATGGGTCAGCCAGCTGCGGATCATCTTC 1036

Db 901 TCCCCACCTTCTGTTGGAGCTCTCTACCATGGGTCAGCCAGCTGCGGATCATCTTC 960

QY 1037 TATATGGCTGCTGTGAAACAAGATGCTGGATGCTTGTGATGCTGCTGCGGCTGAGAGCATGAG 1096

Db 961 TATATGGCTGCTGTGAAACAAGATGCTGGATGCTTGTGATGCTGCTGCGGCTGAGAGCATGAG 1020

QY 1097 ACAATGACACGACAAAGGTCGAGAGACAGTTGGGTTCTACTCTCCGCTCTTCTGGG 1156

Db 1021 ACAATGACACGACAAAGGTCGAGAGACAGTTGGGTTCTACTCTCCGCTCTTCTGGG 1080

QY 1157 GCCATGACGCTGTTGCTTCTCTACCTGCCCCCTCATTTGGGTACATCATGAGTGGCGG 1216

Db 1081 GCCATGACGCTGTTGCTTCTCTACCTGCCCCCTCATTTGGGTACATCATGAGTGGCGG 1140

QY 1217 ATCAAGGACTGCTGGAAGCGCCCACTGAGGACATGCTCTGAGATGCGGAGGAGCGG 1276

Db 1141 ATCAAGGACTGCTGGAAGCGCCCACTGAGGACATGCTCTGAGATGCGGAGGAGCGG 1200

QY 1277 GTTGTACCAATCCATCAGACCGCTACTGCAAGATCCAAAGCTCAACCAATGCGCATC 1336

Db 1201 GTTGTACCAATCCATCAGACCGCTACTGCAAGATCCAAAGCTCAACCAATGCGCATC 1260

QY 1337 AGTGCCTTCAACCTGACCAACCTGCTGCTTGTGGGTTTGGGATCACTGCTGCTCATCAAC 1396

Db 1261 AGTGCCTTCAACCTGACCAACCTGCTGCTTGTGGGTTTGGGATCACTGCTGCTCATCAAC 1320

QY 1397 AACTTACACTCAGTTGTGACCTTTGTCTGCGACACCATTTGTTGAGGTTTCTTCCAC 1456

Db 1321 AACTTACACTCAGTTGTGACCTTTGTCTGCGACACCATTTGTTGAGGTTTCTTCCAC 1380

QY 1457 TCAGGCTGTGGAGTCTCTATGCTGAGTGTTCCTCAACCACTTTTGGAGCGCTGACA 1516

Db 1381 TCAGGCTGTGGAGTCTCTATGCTGAGTGTTCCTCAACCACTTTTGGAGCGCTGACA 1440

QY 1517 GGCCTGCACTCCCTCATCAGTCTGTTGTGCGCTTGTGCTTTCAGCAGCCTTTTTCATGGG 1576

Db 1441 GGCCTGCACTCCCTCATCAGTCTGTTGTGCGCTTGTGCTTTCAGCAGCCTTTTTCATGGG 1500

QY 1577 ATGGTGGAGCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGGCTCTCTGCTATTTCTCA 1636

Db 1501 ATGGTGGAGCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGGCTCTCTGCTATTTCTCA 1560

1637 CTCCTGGGATTCCTGCTGCTTCTTACCTCTTCTATTACCGTCCCGCTCCAGCAGGAG 1696  
 |||||  
 Db 1561 CTCCTGGGATTCCTGCTGCTTCTTACCTCTTCTATTACCGTCCCGCTCCAGCAGGAG 1620  
 |||||

1697 TACGCCGCAATGGGATGGGCCACTGAAGGTGCTTAGCGGCTCTGAGGTGACCGCA 1753  
 |||||  
 Db 1621 TACGCCGCAATGGGATGGGCCACTGAAGGTGCTTAGCGGCTCTGAGGTGACCGCA 1677  
 |||||

BC053747 2507 bp mRNA linear ROD 08-OCT-2003  
 Mus musculus solute carrier family 43, member 1, mRNA (cDNA clone  
 MGC:59491 IMAGE:6330614), complete cds.

ACCESSION BC053747  
 VERSION BC053747.1 GI:31753063  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2507)  
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stepieton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheet, T.E., Brownstein, M.J., Udén, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,  
 Schnerker, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257  
 12477932

2 (bases 1 to 2507)  
 Strausberg, R.  
 Direct Submission  
 Submitted (13-JUN-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
 Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,  
 Ph.D.  
 cDNA Library Preparation: ResGen, Invitrogen Corp  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bcgsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Len Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
 Ness, Pawan Pandoh, Anna-Liisa Parvahu, Parvaneh Saeedi, Jacqueline  
 Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott,  
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 109 Row: b Column: 18  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis.

# FEATURES

## source

1. .2507  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="MGC:59491 IMAGE:6330614"  
 /tissue\_type="Embryo, day 9 mouse (C57BL/6 background)  
 otocysts"  
 /clone\_lib="NIH\_MGC\_130"  
 /lab\_host="DH10B"  
 /note="vector: pCMV-SPORT6.1"  
 1. .2507  
 /gene="Slc43a1"  
 /note="synonyms: PB39, R00504"  
 /db\_xref="LocusID:72401"  
 /db\_xref="MGI:1931352"  
 85..1779  
 /codon\_start=1  
 /product="Slc43a1 protein"  
 /protein\_id="AAH53747.1"  
 /db\_xref="GI:31753064"  
 /db\_xref="LocusID:72401"  
 /translation="MAPTLKQARRRRWMACTAVVNLFPSSAVLLHWSLLIMLKKEG  
 FYSGLCPAENRTNTQDEHQWTSQDOERKMLNGFTIGSFLSATTLPLUGILMDRFG  
 PRRLVGSACFAASCTMALARDTEVLSPLIFLALSLNGFAGICITFTSLTFNPMF  
 GNRSTFMALMIGSYASSAITFGIKLIYDAGVPTVIMFTWSGLACILFNLCAINWP  
 AEAPPAPEVDYTKIKLIGLADHKVTDYFTHVTIVGQRLSKQSPSLSEEGADAFI  
 SSPIDICTSEBTEPKSVPPFKSLCSPIFLMSLATMGMTOQLRVFYMGAMMKILRFIVT  
 GKERETNEQRKVEETVEFYISFGVQMLLCLTCLFLIGYIMDWRNKCDVDATFEGT  
 LNNASFGDARDGASTKTFPRYKRVQKLNAINATLITNILLVFGIACILKHLHQ  
 LLAPVLHTIVRGFFHSACGLYAAVFPSPNHFGLTGQSLTSVAFALLQQLFLPMVVG  
 PLHGDPFWNLGLLILSLFLGFLPLPSLYLYVRSRLQREYATNLVDPQVLNTSKVAT"

## gene

## CDS

## ORIGIN

Query Match 56.7%; Score 1318; DB 10; Length 2507;  
 Best Local Similarity 76.3%; Pred. No. 4.8e-311;  
 Matches 1758; Conservative 0; Mismatches 465; Indels 81; Gaps 8;

QY 66 TCGGGGGCAGCATGGCCCCCAGCGTGCACAGGCGTACCGGAGCGCTGGTGGATGGCT 125  
 |||||  
 Db 74 TCTGGCCACCATGGCTCCACAGCTGAAGAGGGGTACCGAGCGCTGGTGGATGGCT 133  
 |||||

QY 126 GCACGGCTGTGTCGAGAAACCTCTTCTCTCTGTCTGTCTCTCTGGGTGGGGCTCCCTGT 185  
 |||||  
 Db 134 GCACGGCTGTGTCGAGAAACCTCTTCTCTCTGGGTGTCTCTGGGTGGGGCTCCCTGC 193  
 |||||

QY 186 TGATCATCTTGAAGACGAGGGCTTCTATTCCACAGCTGCCAGCTGAGAGCAGCACCA 245  
 |||||  
 Db 194 TGATCATGCTCAAGAGGAGGGCTTCTATTCCAGCTGTGCCAGCTGAGAGCAGCACCA 253  
 |||||

QY 246 ACACACCCAGGATGAGCAGCGCAGGTGGCCAGGCTGTGACACGACGACGACAGATGCTCA 305  
 |||||  
 Db 254 ATACACCCAGATGACAGCATCAGTGGACAACTGTGACACGACGACGAGGAGATGCTCA 313  
 |||||

QY 306 ACCTGGGCTTCAACATTGGTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 365  
 |||||  
 Db 314 ACCTGGGTTTCAACATTGGCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 373  
 |||||

QY 366 TCATGGACCGCTTTGGCCCCCGACCGCTGGCGGTGGTGGCAGTGCCTTCACTGGGT 425  
 |||||  
 Db 374 TCATGGACCGCTTTGGCCCCCGACCGCTTTCACATGGTGGGCGAGTGCCTTTGCCGCAT 433  
 |||||

QY 426 CCTGACCGCTCATGGCCTTGGCCTCCCGGGACGTGGAGAGCTGTGTCTCCGTTGATATTC 485  
 |||||  
 Db 434 CTTGACCTCTAATGGCCTTGGCCTCCAGGGACACTGAAGTTTGTCTCCATGATATTC 493  
 |||||

QY 486 TGGCGCTGTCCCTGAATGGCTTTGGTGGCATCTCCCTAAACGTTCACTTCAGCTGCG 545  
 |||||

Db 494 TGGCACTGTCTTGAATGATTTGCTGGCATCTGCTTAAAGCTTTACCTCACTCACGCTGC 553  
 Qy 546 CCAACATGTTTGGGAACCTGCGCTCCACGTTAATGCGCTCATGATTTGGCTCTTACGCTT 605  
 Db 554 CCAACATGTTTGGGAATTTGCGATCCACATTTCAAGGCTCTCATGATTTGGCTCTTACGCT 613  
 Qy 606 CTTCTGCCATTAAGTTCCAGGAATCAAGCTGATCTACGATGCGGCTGTGGCTTCTGCTG 665  
 Db 614 CTTCTGCCATCAAGTTCCCTTGGGAATCAAGCTGATCTACGATGCGGAGTCCCTTCACTG 673  
 Qy 666 TCAATATGTTCACTGCTGCTGGCTGCGCTGCGCTTATCTTTCTGAATGCGACCTCAACT 725  
 Db 674 TCAATATGTTCACTGCTGCTGGCTGCGCTGCTTATCTTTTGAATGCTGCTCTCAACT 733  
 Qy 726 GGGCCATCAAGCTTTCTGCGCTGCGCTGAGGAAGTCAATTACACGAAGAAGATCAAGCTCA 785  
 Db 734 GGGCTCGAAGAGCTTTCTGCGCTTGGAGAGTGTGACTACACGAAGAAGATCAAACTCA 793  
 Qy 786 GTGGGCTGGCCCTGGACCAAGGTCAGAGTGACCTTCTTACACCCATGTGACCACTCA 845  
 Db 794 TTGGGTTAGCTTTGGACCAAGGTCACAGGTGACCGCTTCTACACCCATGTAAACCATTTG 853  
 Qy 846 TGGGCGAGAGCTCAGCAGAGAGGCCCCAGCTGAGGAGAGGTTCCGATGCTTCACTGT 905  
 Db 854 TGGGTCAGGCGCTGAGTCAAGTCTCCAGGCTGAGGAGGCGCTGACGCGCTTTATTTT 913  
 Qy 906 CACCCAGGATGTTGGGCGACCTCAGAAACCTTCTGAGAGTGTGCTCCCTTACGCA 965  
 Db 914 CATCCCCGAGATATCCCTGTGATCTCAGAGAGACTCTGAAAGTCTGTCCCTTTTCGCA 973  
 Qy 966 AGAGCTCTGCTCCCTCCACTTTCTGTGGAGCTCTCTCACCATGGGCATGACCCAGCTGC 1025  
 Db 974 AGAGCTCTGCTCCCTCCACTTTCTGTGGAGCTTGTCTACCATGGGCATGACCCAGCTTC 1033  
 Qy 1026 GATCATCTTATCATGCTGTGTGAAACAAGATGCTGAGATACCTTTGATGCTGTGGCC 1085  
 Db 1034 GGGTCATCTTATATGGGTGCTATGAACAAGATCTTGGAGTTCAATGTGACTGTGGCA 1093  
 Qy 1086 AGGAGCATGAGCAAAATGACAGCAACAAGGTGGCAGAGCAGTGGTCTACTCTCT 1145  
 Db 1094 AGGAACGTGAGCAAAATGACAGAGAGAGAGAGTGGAGAGAGATGTGAGTTCTACTCTT 1153  
 Qy 1146 CCGCTTTTGGGGCCATGCAAGCTGTTGTGCTTCTCACCTGCCCCCTCATTTGGCTACATCA 1205  
 Db 1154 CCATCTTTGAGTGCAGCTGTTGTGCTTCTCACCTGCCCCCTCATTTGGCTACATCA 1213  
 Qy 1206 TGGACTGGCGGATCAAGGACTGCTGGAGAGCCGCCAACTCAGGGCACTGT----- 1254  
 Db 1214 TGGACTGGCGGATCAAGGACTGTGTGATGCTCCAAACGGAGGGCACCTGAAATGAGATG 1273  
 Qy 1255 -----CCTCGGAGATCCAGGAGCGGTTGCTACCAAAATCCATCAGACCAAGCTACTGCA 1310  
 Db 1274 CTTTCTTTGGAGATGCCAGAGATGGGGCTAGCACCAGTTCTACAGCACGCTACCGCA 1333  
 Qy 1311 AGATCCAAAGCTCACCAGATGCCATCAGTGCCTTTACCCCTGACCAACCTGCTGTGTGG 1370  
 Db 1334 AGGTACAAAGCTCACCAGATGCCATCAATGCCCTTACCCCTGACCAACCTGCTGTGG 1393  
 Qy 1371 GTTTTGGCATCACTGTCTCATCAACAACTTACACCTCAGATTTGTGACTTTGCTCTGC 1430  
 Db 1394 GTTTTGGCATCGCTGCTCATCAAGAACTTACACCTGAGTTGCTGGCTTTGCTCTGC 1453  
 Qy 1431 ACACCATTTTGGAGTCTTCTTCACTCAGCTGCTGGAGTCTCTATGCTGAGTGTTC 1490  
 Db 1454 ATACCATAGTTGCGGTTTCTTCCACTCAGCTGTGGAGTCTCTACGCTGTGTGTTC 1513  
 Qy 1491 CATCCAAACACTTTGGAGGCTGACAGGCTGCTGAGTCCCTCATCAGTGTGTGTGCTCT 1550  
 Db 1514 CGTCCCAATCATTTTGGGACACTGACAGGCTTCTCAGTCTCTCATCAGTGGCTTTGCTC 1573  
 Qy 1551 TGCTTCAGCAGCCACTTTTTCATGGCATGGTGGAGCCCTGAAAGAGAGCCCTTCTGGG 1610  
 Db 1574 TGCTGCAACAGCTACTCTTCTCATGGCCATGGTGGAGCCCTGCAATGGAGATCCCTTCTGG 1633

Qy 1611 TGAATCTGGGCTCTGCTGCTTATTTCTCACTCTCTGGATTCTGTTGCTTCTCTACTCTTCT 1670  
 Db 1634 TGAACCTGGGCTCTCTACTTCTCTGTTCTGGATTCTCTCTACTCTCTACTCTACT 1693  
 Qy 1671 ATTACCGTCCCGGCTCCAGCAGGAGTACGCCGCCAATGGGATGGGCCCACTGAAGTGC 1730  
 Db 1694 ACTACCGGCTCTCGCTGCGAGAGAGATATGCCCAATTTGGTTAGACCCACAGAGGTCG 1753  
 Qy 1731 TTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGAGCTGTGATGACAGCAA 1790  
 Db 1754 TCAATATCTTGAAGTGGCTATACATAGACTCTGAGGCCAAGAGACTTGGAGGACAGCAG 1813  
 Qy 1791 TCAAGGCTCTGAGCAACCAAAAGAGTGCCTCATATGGCTTTTCTACTCTGTAAACATGCACA 1850  
 Db 1814 TCAAGGCTGATTAACCCGAGGAGTGGCTGT---GGCTTTCTACTGTCATCGTGTTC 1870  
 Qy 1851 TAGAGCCATG-CCGCTGAGTATTATAAATACCAAGAGAAAGTCTTATTTTGTAAAGTGC 1909  
 Db 1871 TAGAGCCGGTCTGTGGATTTATAAATACTAAG---AGTTCTATTTTGTAGGACTTG 1927  
 Qy 1910 AAAAGGAGGAAAAA-----CCTTCAAAACGCCC 1942  
 Db 1928 CAAAAAGGAAACAAACAGAACAAACCCCAAAACCAAAACAGCCCTTAAAAAATCTCC 1987  
 Qy 1943 CTAAGTCAACGCTCCATTGACTGAAGACAGTCCCTTATCTTAGAGGGTGTGAGCTTTCTT 2002  
 Db 1988 CCAACCAAGCTATCACTCACTGAAGACAGTCCCGTTCAGTGTATTTGGCCATCGT 2047  
 Qy 2003 CTTCTTGGGTTGGAGGACAGGCTGCTCTTATCTCTTCTAGCGGTCTGCTCTCTG 2062  
 Db 2048 CTTTGAAGAAAGTGTGGGC-----TGCCCTTTCTCCACACTTGGCCTAG 2094  
 Qy 2063 GTACCTTTGGGGGATCGGCAACAGGCTACCCCTGAGGTCCCATGTGCCATGAGTGTG 2122  
 Db 2095 GAGCCTCTGTGGAGTTCCTTCTGAAACTCATGGGATCAGTGAACAGACACACTAAGA 2154  
 Qy 2123 CACAACATCAATGTCTGTGTATGTGTGAATGTGAGAAAAACACAGCCCTCTTTCTAG 2182  
 Db 2155 TTCCAGCCCAAGTGTGCACATATATTTACGCACATGATTTGATCTTTCTCTG 2214  
 Qy 2183 AAGGAAAGGGGCTGAGGTGCGAGCTGTCTCTGGGTTAGGGGTTGGGGGCTCGGCCCTT 2242  
 Db 2215 CAGGAGA-AGGACTGAGGGCTGGCTGTATCCAGGATGGGTGTGGGGCT----- 2264  
 Qy 2243 CCAGGCGCAGGAAGGAGGTTCCCTCTCTGCTGCTGCTGCTTGCAGAGTCTTAGAGAAAT 2302  
 Db 2265 -----GGAGGGGGTTCCTCCCTGATGCTGTTCTTACAGGCTCTTAGAGAAAT 2315  
 Qy 2303 AAAAGGGAAGTGAAGAAAAA 2326  
 Db 2316 AAAAGTGAATGAAGAAAAA 2339

RESULT 11  
 AP001332/c  
 LOCUS  
 DEFINITION Homo sapiens chromosome 11 clone CMB9-114K6 map 11q12, WORKING  
 DRAFT SEQUENCE, 27 unordered pieces.  
 AP001332  
 ACCESSION AP001332.3 GI:9927272  
 VERSION  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 REFERENCE 1 (bases 1 to 136678)  
 AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 TITLE  
 JOURNAL  
 REFERENCE 2 (bases 1 to 136678)  
 AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,







Quality coverage: 4.7 in Q20 bases; agarose-fp  
Quality coverage: 5.8 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 14 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 2043: contig of 2043 bp in length  
\* 2044 2143: gap of 100 bp  
\* 2144 6624: contig of 4481 bp in length  
\* 6625 6724: gap of 100 bp  
\* 6725 14152: contig of 7428 bp in length  
\* 14153 14252: gap of 100 bp  
\* 14253 19511: contig of 5259 bp in length  
\* 19512 19611: gap of 100 bp  
\* 19612 22594: contig of 2983 bp in length  
\* 22595 22694: gap of 100 bp  
\* 22695 27824: contig of 5130 bp in length  
\* 27825 27924: gap of 100 bp  
\* 27925 34952: contig of 7028 bp in length  
\* 34953 35052: gap of 100 bp  
\* 35053 40782: contig of 5730 bp in length  
\* 40783 40882: gap of 100 bp  
\* 40883 48999: contig of 8117 bp in length  
\* 49000 49099: gap of 100 bp  
\* 49100 62883: contig of 13784 bp in length  
\* 62884 62983: gap of 100 bp  
\* 62984 75527: contig of 12544 bp in length  
\* 75528 75628: gap of 100 bp  
\* 75629 87944: contig of 12316 bp in length  
\* 87945 88044: gap of 100 bp  
\* 88045 111029: contig of 23986 bp in length  
\* 111030 111129: gap of 100 bp  
\* 111130 143899: contig of 32770 bp in length.

FEATURES  
source

1. 143899  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-10E21"  
/clone\_lib="RP11-11 Human Male BAC"

misc\_feature

143899  
/note="assembly\_fragment"

misc\_feature

2144..6624  
/note="assembly\_fragment"

misc\_feature

6725..14152  
/note="assembly\_fragment"

misc\_feature

14253..19511  
/note="assembly\_fragment"

misc\_feature

19612..22594  
/note="assembly\_fragment"

misc\_feature

22695..27824  
/note="assembly\_fragment"

misc\_feature

27925..34952  
/note="assembly\_fragment"

misc\_feature

35053..40782  
/note="assembly\_fragment"

misc\_feature

40883..48999  
/note="assembly\_fragment"

misc\_feature

49100..62883  
/note="assembly\_fragment"

misc\_feature

62984..75527  
/note="assembly\_fragment"

misc\_feature

75628..87943  
/note="assembly\_fragment"

misc\_feature

143899  
/note="assembly\_fragment"

misc\_feature 88044..111029  
/note="assembly\_fragment"  
misc\_feature 111130..143899  
/note="assembly\_fragment"

ORIGIN

Query Match 29.4%; Score 684.2; DB 2; Length 143899;  
Best Local Similarity 97.9%; Pred. No. 7.1e-156; Indels 2; Gaps 1;  
Matches 704; Conservative 0; Mismatches 13;  
Qy 1602 CCTTCTGGGTGAATCTGGGCTCTGCTATTCTCACTCTCGGATTCCTGTGCTTCTTCT 1661  
Db 38731 CCCTAAAGGTGAATCTGGGCTCTGCTATTCTCACTCTCGGATTCCTGTGCTTCTTCT 38672  
Qy 1662 ACCTCTTTTATTACCGTCCCGGCTCCAGCAGGATAGCGGCCCAATGGGATGGCCAC 1721  
Db 38671 ACCTCTTTTATTACCGTCCCGGCTCCAGCAGGATAGCGGCCCAATGGGATGGCCAC 38612  
Qy 1722 TGAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1781  
Db 38611 TGAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 38552  
Qy 1782 GACAGGCAATCAAGGCTGAGCAACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTGTA 1841  
Db 38551 GACAGGCAATCAAGGCTGAGCAACCAAAAGGAGTGGCCCATATGGCTTTTCTACCTGTA 38492  
Qy 1842 ACATGCACATAGACCATGGCGCTAGATTTTATAATACCAAGAGAGTCTTATTTTGTGA 1901  
Db 38491 ACATGCACATAGACCATGGCGCTAGATTTTATAATACCAAGAGAGTCTTATTTTGTGA 38432  
Qy 1902 AAGACTGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1961  
Db 38431 AAGACTGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38372  
Qy 1962 GACTGAAGACAGTCCCTATCTAGAGGGTGTAGCTTTCTTCTTCTTCTTCTTCTTCTTCT 2021  
Db 38371 GACTGAAGACAGTCCCTATCTAGAGGGTGTAGCTTTCTTCTTCTTCTTCTTCTTCTTCT 38312  
Qy 2022 ACCAGGTGCTCTTATCTCTTCTAGGGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2081  
Db 38311 ACCAGGTGCTCTTATCTCTTCTAGGGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 38252  
Qy 2082 GCAAAAGGCTACCCCTGAGTCCCATGTGCATGAGTGTGCACAAACATGCAATGTGTCT 2141  
Db 38251 GCAAAAGGCTACCCCTGAGTCCCATGTGCATGAGTGTGCACAAACATGCAATGTGTCT 38194  
Qy 2142 GTGTATGTGAATGTGAGAAACACAGCCCTCTTTCAGAGGAAAGGGCCTGAGGT 2201  
Db 38193 GTGTATGTGAATGTGAGAGACACAGCCCTCTTTCAGAGGAAAGGGCCTGAGGT 38134  
Qy 2202 GCCAGCTGTCTCTGGGTAGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGG 2261  
Db 38133 GCCAGCTGTCTCTGGGTAGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGG 38074  
Qy 2262 TTCCCTCTCTGGTCTGCTGCTTCAAGTCTTACAGGAAATAAAAGGAAAGTGCAGAAA 2320  
Db 38073 TTCCCTCTCTGGTCTGCTGCTTCAAGTCTTACAGGAAATAAAAGGAAAGTGCAGAAA 38015

RESULT 13

AX411543/c  
LOCUS AX411543 143899 bp DNA linear PAT 14-JUN-2002  
DEFINITION Sequence 15 from Patent WO0229059.  
ACCESSION AX411543  
VERSION AX411543.1 GI:21444137  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Sah,D.W.Y., Cate,R.L. and Strittmatter,S.M.  
TITLE Nogo receptor homologs







\* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 15621: contig of 15621 bp in length  
\* 15622: gap of 100 bp  
\* 15721: contig of 100 bp  
\* 15722: contig of 1101 bp in length  
\* 16822: contig of 100 bp  
\* 16823: gap of 100 bp  
\* 16922: gap of 100 bp  
\* 16923: contig of 1864 bp in length  
\* 18787: gap of 100 bp  
\* 18887: contig of 1898 bp in length  
\* 20785: gap of 100 bp  
\* 20786: gap of 100 bp  
\* 20885: contig of 1746 bp in length  
\* 22630: gap of 100 bp  
\* 22631: contig of 2913 bp in length  
\* 25643: gap of 100 bp  
\* 25644: gap of 100 bp  
\* 27574: contig of 2108 bp in length  
\* 27851: gap of 100 bp  
\* 27852: contig of 8082 bp in length  
\* 27952: contig of 100 bp  
\* 108834: gap of 100 bp  
\* 108934: gap of 100 bp  
\* 122115: contig of 13182 bp in length  
\* 122116: gap of 100 bp  
\* 122117: contig of 100 bp  
\* 122118: gap of 100 bp  
\* 122119: contig of 22834 bp in length  
\* 145050: gap of 100 bp  
\* 145150: contig of 17658 bp in length  
\* 162808: gap of 100 bp  
\* 162907: contig of 100 bp  
\* 162908: contig of 2527 bp in length.

FEATURES

Source

1. .165434  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP11-676P16"  
/clone\_lib="RP11-11 Human Male BAC"

misc\_feature

1. .15621  
/note="assembly\_fragment"

clone\_end:SP6

vector\_side:left

15722..16822

/note="assembly\_fragment"

16923..18786

/note="assembly\_fragment"

18887..20784

/note="assembly\_fragment"

20885..22630

/note="assembly\_fragment"

22731..25643

/note="assembly\_fragment"

25744..27851

/note="assembly\_fragment"

27952..108833

/note="assembly\_fragment"

108934..122115

/note="assembly\_fragment"

122116..145049

/note="assembly\_fragment"

145150..162807

/note="assembly\_fragment"

162908..165434

/note="assembly\_fragment"

clone\_end:T7

vector\_side:right

ORIGIN

Query Match 29.3%; Score 682.6; DB 2; Length 165434;  
Best Local Similarity 97.8%; Pred. No. 1.8e-155;  
Matches 703; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

QY 1602 CCTTCTGGGTGAATCTGGGCTCCTGCTATTCTCACTCTGGGATTCCTGTTCCCTTCCCT 1661  
|||  
Db 144291 CCTAAAGGTGAATCTGGGCTCCTGCTATTCTCACTCTGGGATTCCTGTTCCCTTCCCT 144350  
|||  
QY 1662 ACCTCTTCTATTACCGTCCCGGCTCCAGCAGAGTACCCGCAATGGGATGGGCCAC 1721  
|||  
Db 144351 ACCTCTTCTATTACCGTCCCGGCTCCAGCAGAGTACCCGCAATGGGATGGGCCAC 144410  
|||  
QY 1722 TGAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 1781  
|||  
Db 144411 TGAAGGTGCTTAGCGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGACCTGGAT 144470  
|||  
QY 1782 GACAGCAATCAAGGCTCTGAGCAACCAAAAGGAGTCCCATATGCTTTCTTACCTGTA 1841  
|||  
Db 144471 GACAGCAATCAAGGCTCTGAGCAACCAAAAGGAGTCCCATATGCTTTCTTACCTGTA 144530  
|||  
QY 1842 ACATGCACATAGAGCCATGCGCTAGATTATAAATACCAAGAGAAATGTTCTATTTTGTGA 1901  
|||  
Db 144531 ACATGCACATAGAGCCATGCGCTAGATTATAAATACCAAGAGAAATGTTCTATTTTGTGA 144590  
|||  
QY 1902 AGACTGCAAAAAGGAGGAAAAAACCTTCAAAAAGGCGCCCTAAGTCAACGCTCCATT 1961  
|||  
Db 144591 AGACTGCAAAAAGGAGGAAAAAACCTTCAAAAAGGCGCCCTAAGTCAACGCTCCATT 144650  
|||  
QY 1962 GACTGAAGACAGTCCCTATCTAGAGGGGTGAGCTTTCTTCTCTCTGGTTGGAGGAG 2021  
|||  
Db 144651 GACTGAAGACAGTCCCTATCTAGAGGGGTGAGCTTTCTTCTCTCTGGTTGGAGGAG 144710  
|||  
QY 2022 ACCAGGGTGCCTTTATCTCTTCTAGCGGTCTGCTCTCTGGTACTCTTTGGGGGATCG 2081  
|||  
Db 144711 ACCAGGGTGCCTTTATCTCTTCTAGCGGTCTGCTCTCTGGTACTCTTTGGGGGATCG 144770  
|||  
QY 2082 GCAAAACGGCTACCCCTGAGGTCCCATGTGCCATGATGTGCACACATGCTATGTCT 2141  
|||  
Db 144771 GCAAAACGGCTACCCCTGAGGTCCCATGTGCCATGATGTGCACACATGCTATGTCT 144828  
|||  
QY 2142 GTGTATGTGTAATGTGAGAAAAACACAGCCCTCTCTTTCAGAAAGGAAAGGGGCTCAGGT 2201  
|||  
Db 144829 GTGTATGTGTAATGTGAGAGAGACACAGCCCTCTCTTTCAGAAAGGAAAGGGGCTCAGGT 144888  
|||  
QY 2202 GCCAGCTGTCTCTGGTTAGGGGTTGGGGTTCGGGCTCGGCCCTTCCAGGGCCAGGAGGAGG 2261  
|||  
Db 144889 GCCAGCTGTCTCTGGTTAGGGGTTGGGGTTCGGGCTCGGCCCTTCCAGGGCCAGGAGGAGG 144948  
|||  
QY 2262 TTCCCTCTCTGCTGCTGCTGCTTGCAGTCTTAGAGGAATATAAAGGGAAGTGAGAAA 2320  
|||  
Db 144949 TTCCCTCTCTGCTGCTGCTGCTTGCAGTCTTAGAGGAATATAAAGGGAAGTGAGAGA 145007  
|||

Search completed: April 10, 2004, 17:32:54

Job time : 9401.6 secs

**THIS PAGE BLANK (USPTO)**

[illegible]

Qy	1	CCGGGCTGAGGGGGCAACGGGTTCGAGGTGCAAGCCTGGTGCCTGCCCGAGCCCTGC	60
Db	1	CCGGGCTGAGGGGGCAACGGGTTCGAGGTGCAAGCCTGGTGCCTGCCCGAGCCCTGC	60



Db 2221 AGGGGTGGGGGTCGGCCCTTCCAGGGCCAGAAAGGCAAGGTTCCTCTCTGTGTGTCT 2280  
Qy 2281 GCTTCAAGTCTTAGAGGAATAAAAAAGGAAGTGAAGAAAAA 2326  
Db 2281 GCTTCAAGTCTTAGAGGAATAAAAAAGGAAGTGAAGAAAAA 2326

RESULT 2

US-10-205-823-330  
; Sequence 330, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Womsey, Angela M.  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Anderson, Dustin  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
; FILE REFERENCE: MRI-044  
; CURRENT APPLICATION NUMBER: US/10/205,823  
; CURRENT FILING DATE: 2002-07-25  
; PRIOR APPLICATION NUMBER: 60/307,982  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: 60/314,356  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/325,020  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: 60/341,746  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/362,158  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 455  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 330  
; LENGTH: 2326  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-205-823-330

Query Match 100.0%; Score 2326; DB 14; Length 2326;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CCGGGGCTGGAGGGGGCAAGCGGGTTCGAGGTGCAAGAGCTGTGCGCCGAGCCCTGC 60  
Db 1 CCGGGGCTGGAGGGGGCAAGCGGGTTCGAGGTGCAAGAGCTGTGCGCCGAGCCCTGC 60  
Qy 61 GGAGCTCGGGGGCAGCATGGCCCCCAGCTGCAACAGGGGTACCGAGGCGCTGTGTGGAT 120  
Db 61 GGAGCTCGGGGGCAGCATGGCCCCCAGCTGCAACAGGGGTACCGAGGCGCTGTGTGGAT 120  
Qy 121 GGCTCGACGGCTGTGTGAGAACTCTTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 180  
Db 121 GGCTCGACGGCTGTGTGAGAACTCTTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 180  
Qy 181 CTTGTTGATCATCTCAAGAACGAGGGCTTCTATTCCAGCAGCTGCCCCAGCTGAGAGCAG 240  
Db 181 CTTGTTGATCATCTCAAGAACGAGGGCTTCTATTCCAGCAGCTGCCCCAGCTGAGAGCAG 240  
Qy 241 CACCAACACACCCAGGATGAGCAGCGAGGTGGCCAGGCTGTGACAGCAGGACGAGAT 300  
Db 241 CACCAACACACCCAGGATGAGCAGCGAGGTGGCCAGGCTGTGACAGCAGGACGAGAT 300  
Qy 301 GCTCAACCTGGGCTTCCCAATTGGTTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360  
Db 301 GCTCAACCTGGGCTTCCCAATTGGTTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360

Db 301 GCTCAACCTGGGCTTCCCAATTGGTTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360  
Qy 361 GATCTCTCATGGACCGCTTTTGGCCCCCGACCGCTGGGGCTGGTTGGCAGTGCCTTGTCTCAC 420  
Db 361 GATCTCTCATGGACCGCTTTTGGCCCCCGACCGCTGGGGCTGGTTGGCAGTGCCTTGTCTCAC 420  
Qy 421 TGGGTCTGTCACCCCTCATNGGCCCTTGGCCCTCCCGGGAAGCTGTGTCTCCGTGTGAT 480  
Db 421 TGGGTCTGTCACCCCTCATNGGCCCTCCCGGGAAGCTGTGTCTCCGTGTGAT 480  
Qy 481 ATTCTGGGCGCTGTCCCTGAATGGCTTTGGTGGCATCTGCCTAAAGCTTCACTTCACTCAC 540  
Db 481 ATTCTGGGCGCTGTCCCTGAATGGCTTTGGTGGCATCTGCCTAAAGCTTCACTTCACTCAC 540  
Qy 541 GCTGCCCAACATGTTTGGGAACCTGGGCTCCAGCTTAAATGGCCCTCATGATTGGCTCTTA 600  
Db 541 GCTGCCCAACATGTTTGGGAACCTGGGCTCCAGCTTAAATGGCCCTCATGATTGGCTCTTA 600  
Qy 601 CGCTCTTCTGTCATTTACGTTTCCAGGAATCAAGCTGTATACGATGCCGGTGTGGCCCTT 660  
Db 601 CGCTCTTCTGTCATTTACGTTTCCAGGAATCAAGCTGTATACGATGCCGGTGTGGCCCTT 660  
Qy 661 CGTGTCTCATGTTTCACTGTCTGGCCCTGGCCCTGCCTTATCTTCTGAATGCAACCT 720  
Db 661 CGTGTCTCATGTTTCACTGTCTGGCCCTGGCCCTGCCTTATCTTCTGAATGCAACCT 720  
Qy 721 CAACTGGCCCATCGAAGCTTTCCTGCCCCCTGAGGAAGTCAATTACACGAAGAAGATCAA 780  
Db 721 CAACTGGCCCATCGAAGCTTTCCTGCCCCCTGAGGAAGTCAATTACACGAAGAAGATCAA 780  
Qy 781 GCTGAGTGGGCTGGCCCTGGACCAAGGTGACAGGTGACCTTCTTACACCCATGTGAC 840  
Db 781 GCTGAGTGGGCTGGCCCTGGACCAAGGTGACAGGTGACCTTCTTACACCCATGTGAC 840  
Qy 841 CACCATGGCCAGAGCTCAGCCAGAGGCCGCCAGCTGGAGGAGCGTTCGGATGCCCTT 900  
Db 841 CACCATGGCCAGAGCTCAGCCAGAGGCCGCCAGCTGGAGGAGCGTTCGGATGCCCTT 900  
Qy 901 CATGTCAACCCAGGATGTTTCGGGGCACCTCAGAAAACCTTCTGAGAGGTCTGTGCCCTT 960  
Db 901 CATGTCAACCCAGGATGTTTCGGGGCACCTCAGAAAACCTTCTGAGAGGTCTGTGCCCTT 960  
Qy 961 ACACAAGAGCTCTGCTCCCCCACTTCTGTGGAGCTTCTCACCATGGGATGACCCCA 1020  
Db 961 ACACAAGAGCTCTGCTCCCCCACTTCTGTGGAGCTTCTCACCATGGGATGACCCCA 1020  
Qy 1021 GCTGGGATCATCTTCTACATGGCTGTGTGAAACAGATGCTGGAGTACCTTGTGTGATGG 1080  
Db 1021 GCTGGGATCATCTTCTACATGGCTGTGTGAAACAGATGCTGGAGTACCTTGTGTGATGG 1080  
Qy 1081 TGGCCAGGAGCATGAGACAAATGAACAGCAACAAAGGTGGCAGAGACAGTTGGGTTCTA 1140  
Db 1081 TGGCCAGGAGCATGAGACAAATGAACAGCAACAAAGGTGGCAGAGACAGTTGGGTTCTA 1140  
Qy 1141 CTCCTCCGTCTTTCGGGGCCATGACGTGTGTGCTTCTCACCCTGCCCTCATTTGGCTA 1200  
Db 1141 CTCCTCCGTCTTTCGGGGCCATGACGTGTGTGCTTCTCACCCTGCCCTCATTTGGCTA 1200  
Qy 1201 CATCATGGAGCTGCGGATCAAGGACTGCGTGGACGCCCACTCAGGGGACCTGTCTCGG 1260  
Db 1201 CATCATGGAGCTGCGGATCAAGGACTGCGTGGACGCCCACTCAGGGGACCTGTCTCGG 1260  
Qy 1261 AGATGTCAGGGAGCGGGTGTCTACCAATCCATCAGACCACTACTGCAAGATCCAAA 1320  
Db 1261 AGATGTCAGGGAGCGGGTGTCTACCAATCCATCAGACCACTACTGCAAGATCCAAA 1320  
Qy 1321 GCTCAACCAATGCCATCAGTGCCTTCACTGACCAACCTGTGTGTGGGTTTGGCAT 1380  
Db 1321 GCTCAACCAATGCCATCAGTGCCTTCACTGACCAACCTGTGTGTGGGTTTGGCAT 1380  
Qy 1381 CACTGTCTCATCAACAACTTACACCTCAAGTTGTGACCTTGTCTGTCGACACCATTTGT 1440  
Db 1381 CACTGTCTCATCAACAACTTACACCTCAAGTTGTGACCTTGTCTGTCGACACCATTTGT 1440







```
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 129
; LENGTH: 2284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1768)
US-10-262-839-129

Query Match      92.4%; Score 2150; DB 12; Length 2284;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 10; Indels 47; Gaps 2;

Qy      31 AGGTGCAAGCCTGTGTCCTGCGGAGCCTGCGGAGCTCGGGGCGCAGCATGGCCCCCAGCGCT 90
Db      1 AGGTGCAAGCCTGTGTCCTGCGGAGCCTGCGGAGCTCGGGGCGCAGCATGGCCCCCAGCGCT 60

Qy      91 GCAACAGGCGTACCGGAGCGCTGTGATGCGCTGCAAGGCTGTGCTGGAGAACCTTTT 150
Db      61 GCAACAGGCGTACCGGAGCGCTGTGATGCGCTGCAAGGCTGTGCTGGAGAACCTTTT 120

Qy      151 CTTCTCTGTGTACTTCTGGGCTGGGCTCCCTGTGTGATCATCTTGTGAGAACGAGGCTT 210
Db      121 CTTCTCTGTGTACTTCTGGGCTGGGCTCCCTGTGTGATCATCTTGTGAGAACGAGGCTT 180

Qy      211 CTTATCCAGCAGTGCC----- 227
Db      181 CTTATCCAGCAGTGCCAGCTGTTCTTGGTGTATGTGCTGGGCGCTCCCTTCCCCCTC 240

Qy      228 --CAGCTGAGAGCAGCAACCAACCAAGATGAGCAGCGCAGGTCGCGAGGCTGTGA 285
Db      241 CTCAGCTGAGAGCAGCAACCAACCAAGATGAGCAGCGCAGGTCGCGAGGCTGTGA 300

Qy      286 CCAGCAGAGCAGAGTCTCAACTCGGCTTCAACATGCTTCTTCTGCTCAGCGCCAC 345
Db      301 CCAGCAGAGCAGAGTCTCAACTCGGCTTCAACATGCTTCTTCTGCTCAGCGCCAC 360

Qy      346 CACCCTGCCACTGGGGATCCTCATGAGACCGCTTTGGCCCCCGACCCGTCGCGCTGGTTGG 405
Db      361 CACCCTGCCACTGGGGATCCTCATGAGACCGCTTTGGCCCCCGACCCGTCGCGCTGGTTGG 420

Qy      406 CAGTGCCTGCTTCACTGCGTCTGCAACCTCATAGCCCTTGGCCTCCCGGGACGTGGAAGC 465
Db      421 CAGTGCCTGCTTCACTGCGTCTGCAACCTCATAGCCCTTGGCCTCCCGGGACGTGGAAGC 480

Qy      466 TCTGTCTCGTTGATATTCCTGCGCTGTCCTGAATGGCTTGGTGGCATCTGCTTAAC 525
Db      481 TCTGTCTCGTTGATATTCCTGCGCTGTCCTGAATGGCTTGGTGGCATCTGCTTAAC 540

Qy      526 GTTCACCTTCACTCAGCGTGCCCAACATGTTTGGGAACCTTGGCCTCCAGCTTAATGGCCCT 585
Db      541 GTTCACCTTCACTCAGCGTGCCCAACATGTTTGGGAACCTTGGCCTCCAGCTTAATGGCCCT 600

Qy      586 CATGATTTGGCTTTAGCGCTCTTCTGGCAATTAAGTTCCCGAGGAATCAAGCTGATCTAGA 645
Db      601 CATGATTTGGCTTTAGCGCTCTTCTGGCAATTAAGTTCCCGAGGAATCAAGCTGATCTAGA 660

Qy      646 TSCCGGTGGCTTTCGTGGTTCATCATGTTACCTGCTGTCGCTGGCTGCTTTATCTT 705
Db      661 TSCCGGTGGCTTTCGTGGTTCATCATGTTACCTGCTGTCGCTGGCTGCTTTATCTT 720

Qy      706 TCTGAAGTGCACCTCAACTGGGCCCATCGAAGCCTTTCTGCCCCCTGAGGAAGTCAATTA 765
```

```
Db      721 TCTGAAGTGCACCTCAACTGGGCCCATCGAAGCCTTTCTTCCCTCGAGGAAGTCAATTA 780
Qy      766 CACGAAGAAGATCAAGCTGAGTGGGCTGGGCTGGACCAACAAGTGACAGGTGACCTCTT 825
Db      781 CACGAAGAAGATCAAGCTGAGTGGGCTGGGCTGGACCAACAAGTGACAGGTGACCTCTT 840
Qy      826 CTACACCATGTGTACCAACCATGGCCAGAGGCTCAGCCAGAGGCCCCAGCCTGGAGGA 885
Db      841 CTACACCATGTGTACCAACCATGGCCAGAGGCTCAGCCAGAGGCCCCAGCCTGGAGGA 900
Qy      886 CGGTTCCGATGCTTCTATGTACCCAGAGTGTTCGGGGCACCTTCAGAAAACTTCTCTGA 945
Db      901 CGGTTCCGATGCTTCTATGTACCCAGAGTGTTCGGGGCACCTTCAGAAAACTTCTCTGA 960
Qy      946 GAGTCTGTCCCTTACGCAAGAGCCTTCTGCTCCCACTTTCTCTGTGGAGCCTCTCTAC 1005
Db      961 GAGTCTGTCCCTTACGCAAGAGCCTTCTGCTCCCACTTTCTCTGTGGAGCCTCTCTAC 1020
Qy      1006 CATGGGATGACCCAGCTGCGGATCATCTTCTACATGGCTGCTGTGAACAGATGCTGGA 1065
Db      1021 CATGGGATGACCCAGCTGCGGATCATCTTCTACATGGCTGCTGTGAACAGATGCTGGA 1080
Qy      1066 GTACCTTTGTGATGCTGGTGGCAGGAGTGAACAATAAAGCAAGCAAAAGGTGGCAGA 1125
Db      1081 GTACCTTTGTGATGCTGGTGGCAGGAGTGAACAATAAAGCAAGCAAAAGGTGGCAGA 1140
Qy      1126 GACAGTTGGGTTTCTACTCTCCGCTTTCGGGGCCATGAGAGCTGTGTGCTTCTCACTG 1185
Db      1141 GACAGTTGGGTTTCTACTCTCCGCTTTCGGGGCCATGAGAGCTGTGTGCTTCTCACTG 1200
Qy      1186 CCCCCTCATTTGGCTTACATCATGAGTGGCGATCAAGAGCTGCGTGAAGGCCCAACTCA 1245
Db      1201 CCCCCTCATTTGGCTTACATCATGAGTGGCGATCAAGAGCTGCGTGAAGGCCCAACTCA 1260
Qy      1246 GGGCAGCTGTCTCGGAGATGCCAGGAGCGGGTTGCTTACCAATCCATCAGACCAACGCTA 1305
Db      1261 GGGCAGCTGTCTCGGAGATGCCAGGAGCGGGTTGCTTACCAATCCATCAGACCAACGCTA 1320
Qy      1306 CTGCAAGATCCAAAGCTCACCAGTGCATGAGTGCCTTCCCTGACCAACCTGCTGCT 1365
Db      1321 CTGCAAGATCCAAAGCTCACCAGTGCATGAGTGCCTTCCCTGACCAACCTGCTGCT 1380
Qy      1366 TGTGGGTTTTGGCATCACCTGTCTCATCAACAACTTACACCTCCAGTTTGTGACCTTGT 1425
Db      1381 TGTGGGTTTTGGCATCACCTGTCTCATCAACAACTTACACCTCCAGTTTGTGACCTTGT 1440
Qy      1426 CCTGCACACCAATTTTGGAGTTTCTTCCACTCAGCTGCTGGGAGTCTCTATGCTGCACT 1485
Db      1441 CCTGCACACCAATTTTGGAGTTTCTTCCACTCAGCTGCTGGGAGTCTCTATGCTGCACT 1500
Qy      1486 GTTCCCATCCAAACCACTTTGGGAGCTGACAGGCTGCTGCTCCCTCATCTGCTGCTGTT 1545
Db      1501 GTTCCCATCCAAACCACTTTGGGAGCTGACAGGCTGCTGCTCCCTCATCTGCTGCTGTT 1560
Qy      1546 CGCTTGTCTCAGCAGCCACTTTTTCATGGCGATGGTGGAGCCCTGAAAGAGAGAGCCCTT 1605
Db      1561 CGCTTGTCTCAGCAGCCACTTTTTCATGGCGATGGTGGAGCCCTGAAAGAGAGAGCCCTT 1620
Qy      1606 CTGGGTGAATCTGGGCTCTCTGCTATTTCTCATCTCTGGGATTTCTGTTGCTTCTTACCT 1665
Db      1621 CTGGGTGAATCTGGGCTCTCTGCTATTTCTCATCTCTGGGATTTCTGTTGCTTCTTACCT 1680
Qy      1666 CTTCTATTACCGTCCCGGCTCCAGCAGGAGTACGCCGCAATGGGATGGGCCCACTGAA 1725
Db      1681 CTTCTATTACCGTCCCGGCTCCAGCAGGAGTACGCCGCAATGGGATGGGCCCACTGAA 1740
Qy      1726 GGTGCTTAGCGGCTCTGAGGTGACCGATAGACTTCTCAGACCAAGGACCTGATGACA 1785
Db      1741 GGTGCTTAGCGGCTCTGAGGTGACCGATAGACTTCTCAGACCAAGGACCTGATGACA 1800
Qy      1786 GGCATCAAGGCTTGAGCAACCAAGAGGTGCGCCCATATGGCTTTTCTACTGTAACT 1845
```



Db 537 -----TCAAGCTGATCTAGCATGCCGTTGGCCTT 567  
Qy 661 CGTGTGTCATGTTTCACTGCTGTGCGCTGCGCTTATCTTCTTGAACGTGACCCCT 720  
Db 568 CGTGTGTCATGTTTCACTGCTGTGCGCTGCGCTTATCTTCTTGAACGTGACCCCT 627  
Qy 721 CAACCTGGCCCATCGAAGCCTTCTTCCCTGCGCTGAGGAAGTCAATTACACGAAGAATCAAA 780  
Db 628 CAACCTGGCCCATCGAAGCCTTCTTCCCTGCGCTGAGGAAGTCAATTACACGAAGAATCAAA 687  
Qy 781 GCTGAGTGGGCTGGCCCTGGACCAACAAGGTGACAGGTGACCTTCTTACACCCATGTGAC 840  
Db 688 GCTGAGTGGGCTGGCCCTGGACCAACAAGGTGACAGGTGACCTTCTTACACCCATGTGAC 747  
Qy 841 CACCATGGGCGAGGCTCAGCCAGAGGCCGCCCTGAGGAGCGGTTCGGATGCGCTT 900  
Db 748 CACCATGGGCGAGGCTCAGCCAGAGGCCGCCCTGAGGAGCGGTTCGGATGCGCTT 807  
Qy 901 CATGTCAACCCAGGATGTTTCGGGGCACCTTCAGAAACCTTCTGAGAGGTCTGTCCCTT 960  
Db 808 CATGTCAACCCAGGATGTTTCGGGGCACCTTCAGAAACCTTCTGAGAGGTCTGTCCCTT 867  
Qy 961 ACGCAAGAGCCTCTGCTCCCCACCTTCTGTGGAGCCTCTCACCATGGGCATGACCCA 1020  
Db 868 ACGCAAGAGCCTCTGCTCCCCACCTTCTGTGGAGCCTCTCACCATGGGCATGACCCA 927  
Qy 1021 GCTGCGGATCATCTTCTACATGGCTGTGTGAACAAGATGCTGGAGTACCTTGTGACTGG 1080  
Db 928 GCTGCGGATCATCTTCTACATGGCTGTGTGAACAAGATGCTGGAGTACCTTGTGACTGG 987  
Qy 1081 TGGCCAGGAGCATGAGCAAAATGAACAGCAAAAGGTGGCAGAGACAGTTGGTCTCTA 1140  
Db 988 TGGCCAGGAGCATGAGCAAAATGAACAGCAAAAGGTGGCAGAGACAGTTGGTCTCTA 1047  
Qy 1141 CTCTCTCCGTCTTTCGGGGCCATGCAGCTGTGTGSCCTTCTCACTGCCCCCTCATTTGGCTA 1200  
Db 1048 CTCTCTCCGTCTTTCGGGGCCATGCAGCTGTGTGSCCTTCTCACTGCCCCCTCATTTGGCTA 1107  
Qy 1201 CATCATGAGCTGCGGATCAAGGACTGTGTGAGCGCCCACTCAGGCGACTGTCTCGG 1260  
Db 1108 CATCATGAGCTGCGGATCAAGGACTGTGTGAGCGCCCACTCAGGCGACTGTCTCGG 1167  
Qy 1261 AGATGCCAGGAGCGGGTGTCTACAAATCCATCAGACACGCTACTGCAAGATCCAAA 1320  
Db 1168 AGATGCCAGGAGCGGGTGTCTACAAATCCATCAGACACGCTACTGCAAGATCCAAA 1227  
Qy 1321 GCTCAACCAATGCCATCAGTGCCTTCACTGACCAACCTGCTGTGTGGGTTTTGGCAT 1380  
Db 1228 GCTCAACCAATGCCATCAGTGCCTTCACTGACCAACCTGCTGTGTGGGTTTTGGCAT 1287  
Qy 1381 CACCTGTCTCATCAACACTTACACTCCAGTTTGTGACCTTGTGCTGCGACACCATTTGT 1440  
Db 1288 CACCTGTCTCATCAACACTTACACTCCAGTTTGTGACCTTGTGCTGCGACACCATTTGT 1347  
Qy 1441 TCGAGGTTCTTCCACTCAGCCTGTGGGAGTCTATGCTGCGAGTGTTCCTCAACCA 1500  
Db 1348 TCGAGGTTCTTCCACTCAGCCTGTGGGAGTCTATGCTGCGAGTGTTCCTCAACCA 1407  
Qy 1501 CTTTGGGAGCGTGAAGGCGCTCAGTCCCTCATCATAGTGTGTGCTTGTGCTTTCAGCA 1560  
Db 1408 CTTTGGGAGCGTGAAGGCGCTCAGTCCCTCATCATAGTGTGTGCTTGTGCTTTCAGCA 1467  
Qy 1561 GCCACTTTTCATGGGATGTTGGGACCCCTGAAAGGAGGCGCTTCTGGGTGAATCTGGG 1620  
Db 1468 GCCACTTTTCATGGGATGTTGGGACCCCTGAAAGGAGGCGCTTCTGGGTGAATCTGGG 1527  
Qy 1621 CTTCTGCTATTTCTCACTCCTGGGATTTCTGTTGCTTCTTCACTCTTCTATTACCGTGC 1680  
Db 1528 CTTCTGCTATTTCTCACTCCTGGGATTTCTGTTGCTTCTTCACTCTTCTATTACCGTGC 1587  
Qy 1681 CCGGCTCAGCAGGATGACGCGCCCAATGGGATGGGCGCCACTGAAAGGTGCTTACGGGCTC 1740

Db 1588 CCGGCTCAGCAGGAGTACGCGCCCAATGGGATGGGCCCACTGAAGGTGCTTACGGGCTC 1647  
Qy 1741 TGAGGTGACCGCATAGACTTCTCAGACCAAGGGACCTGGATGACAGGCAATCAAGGCGCTG 1800  
Db 1648 TGAGGTGACCGCATAGACTTCTCAGACCAAGGGACCTGGATGACAGGCAATCAAGGCGCTG 1707  
Qy 1801 AGCAACCAAAAGAGTGCCTCCCATATGGCTTTTCTPACCTGTAAACATGCACATAGAGCCATG 1860  
Db 1708 AGCAACCAAAAGAGTGCCTCCCATATGGCTTTTCTPACCTGTAAACATGCACATAGAGCCATG 1767  
Qy 1861 GCGCTAGATTATATAATACCAAGAGAGTCTATTTTGTAAAGACTGCAAAAAGGAGGA 1920  
Db 1768 GCGCTAGATTATATAATACCAAGAGAGTCTATTTTGTAAAGACTGCAAAAAGGAGGA 1827  
Qy 1921 AAAAAACCTTCAAAAAAGCGCCCTTAAGTCAACGCTCCATTTGACTGAAGCAGTCCCTAT 1980  
Db 1828 AAAAAACCTTCAAAAAAGCGCCCTTAAGTCAACGCTCCATTTGACTGAAGCAGTCCCTAT 1887  
Qy 1981 CCTAGAGGGTGTAGCTTTTCTTCTCTTGGTTGGAGGAGACCGAGGTGCTCTTATCT 2040  
Db 1888 CCTAGAGGGTGTAGCTTTTCTTCTCTTGGTTGGAGGAGACCGAGGTGCTCTTATCT 1947  
Qy 2041 CTTTCTAGCGTCTGCTCTCTGTAACCTTCTTGGGGGATCGGCAAAAGGCTACCCCTGA 2100  
Db 1948 CTTTCTAGCGTCTGCTCTCTGTAACCTTCTTGGGGGATCGGCAAAAGGCTACCCCTGA 2007  
Qy 2101 GGTCCCATGTGCCATGAGTGTGCAACAATGCTGTCTGTGTATGTGTGAATGTGAG 2160  
Db 2008 GGTCCCATGTGCCATGAGTGTGCAACAATGCTGTGTATGTGTGAATGTGAG 2067  
Qy 2161 AAAAAACAGCCCTCTTTCAGAGGAAAGGGGCTCAGGTGCCAGCTGTCTCTGGGTT 2220  
Db 2068 AAAAAACAGCCCTCTTTCAGAGGAAAGGGGCTCAGGTGCCAGCTGTCTCTGGGTT 2127  
Qy 2221 AGGGGTTGGGGTGGCGCCCTTCCAGGCGCAGGAAGGAGGTTCCTCTCTGGTGTGCT 2280  
Db 2128 AGGGGTTGGGGTGGCGCCCTTCCAGGCGCAGGAAGGAGGTTCCTCTCTGGTGTGCT 2187  
Qy 2281 GCTTGAAGCTTATAGAGGAATAAAGGGAAGTGAAGAAAAA 2326  
Db 2188 GCTTGAAGCTTATAGAGGAATAAAGGGAAGTGAAGAAAAA 2233

RESULT 6  
US-10-262-839-127  
; Sequence 127, Application US/10262839  
; Publication No. US20040038877A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook, John,  
; APPLICANT: Anderson, David W.,  
; APPLICANT: Boldog, Ferenc,  
; APPLICANT: Burgess, Catherine,  
; APPLICANT: Catterton, Elina,  
; APPLICANT: Edinger, Shlomit,  
; APPLICANT: Ellerman, Karen,  
; APPLICANT: Gerlach, Valerie,  
; APPLICANT: Gorman, Linda,  
; APPLICANT: Guo, Xiaojia,  
; APPLICANT: Ji, Weizhen,  
; APPLICANT: Kekuda, Rameah,  
; APPLICANT: Leach, Martin,  
; APPLICANT: Li, Li,  
; APPLICANT: Miller, Charles,  
; APPLICANT: Patturajan, Meera,  
; APPLICANT: Reiger, Daniel,  
; APPLICANT: Rothenberg, Mark,  
; APPLICANT: Shimkets, Richard,  
; APPLICANT: Smithson, Glennnda,  
; APPLICANT: Spytek, Kimberly,  
; APPLICANT: Taupier, Raymond, jr.,  
; APPLICANT: Vernet, Corine,  
; APPLICANT: Voss, Edward,  
; APPLICANT: Zerhusen, Brian,

; APPLICANT: Zhong, Mei  
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 ; FILE REFERENCE: 21402-462A  
 ; CURRENT APPLICATION NUMBER: US/10/262,839  
 ; CURRENT FILING DATE: 2002-10-01  
 ; PRIOR APPLICATION NUMBER: 60/326,483  
 ; PRIOR FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: 60/327,917  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/328,029  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/328,056  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,101  
 ; PRIOR FILING DATE: 2002-05-16  
 ; PRIOR APPLICATION NUMBER: 60/371,972  
 ; PRIOR FILING DATE: 2002-04-12  
 ; PRIOR APPLICATION NUMBER: 60/327,342  
 ; PRIOR FILING DATE: 2001-10-05  
 ; PRIOR APPLICATION NUMBER: 60/328,044  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/328,849  
 ; PRIOR FILING DATE: 2001-10-12  
 ; PRIOR APPLICATION NUMBER: 60/374,738  
 ; PRIOR FILING DATE: 2002-04-23  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 367  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO 127  
 ; LENGTH: 2105  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (47)..(1588)  
 ; US-10-262-839-127

Query Match 82.3%; Score 1914; DB 12; Length 2105;  
 Best Local Similarity 93.4%; Pred. No. 0;  
 Matches 2088; Conservative 0; Mismatches 10; Indels 138; Gaps 3;  
  
 QY 31 AGGTGCAAAAGCGTGTGCCCCGAGCCCTCGGGAGCTCGGGGCGAGCATGGCCCCACGCT 90  
 DB 1 AGGTGCAAAAGCGTGTGCCCCGAGCCCTCGGGAGCTCGGGGCGAGCATGGCCCCACGCT 60  
  
 QY 91 GCACAGGCGTACCGAGGCGCTGGTGGATGGCTGCAAGGCTGCTGGAGAACCTCTT 150  
 DB 61 GCACAGGCGTACCGAGGCGCTGGTGGATGGCTGCAAGGCTGCTGGAGAACCTCTT 120  
  
 QY 151 CTTCTCTGCTGTACTCTTGGGCTGGGCTCCCTGTTGATCTTCAAGAACGAGGGCTT 210  
 DB 121 CTTCTCTGCTGTACTCTTGGGCTGGGCTCCCTGTTGATCTTCAAGAACGAGGGCTT 180  
  
 QY 211 CTATTCAGACAGTGGCCCCAGCTGAGAGCAGCAACACACCCAGGATGAGCAGGCGAG 270  
 DB 181 CTATTCAGACAGTGGCCCCAGCTGAGAGCAGCAACACACCCAGGATGAGCAGGCGAG 240  
  
 QY 271 GTGGCCAGGCTGTGACGAGCAGAGATGCTCAACCTGGGGCTTCAACCATTTGCTTCT 330  
 DB 241 GTGGC----- 245  
  
 QY 331 CGTGCTCAGCGCCACCACTGCGCATCGGGATCCTCATGGACCGCTTGGCCCCCGACC 390  
 DB 246 ----- 245  
  
 QY 391 CGTGCGGCTGGTTGGAGTGCCTGCTTCACTGGCTCTGCACCCCTCATGGCCCTGGCCTC 450  
 DB 246 -----CTTGGCTTCACTGGCTCTGCACCCCTCATGGCCCTGGCCTC 285  
  
 QY 451 CCGGGACGTGGAAGCTCTGTCTCCGTTGATATTCCTGGGCGCTGCTCCCTGAATGGCTTTGG 510  
 DB 286 CCGGGACGTGGAAGCTCTGTCTCCGTTGATATTCCTGGGCGCTGCTCCCTGAATGGCTTTGG 345

QY 511 TGGCATCTCGCTAAACGTTTCACTTCACTCAGCGTGCACAAACATGTTTGGGAACCTGGCGTC 570  
 DB TGGCATCTCGCTAAACGTTTCACTTCACTCAGCGTGCACAAACATGTTTGGGAACCTGGCGTC 405  
  
 QY 571 CACGTTAATGGCCCTCATGATTGGCTTTTACGCTCTTTCGCCATTAAGTTCACGAAGT 630  
 DB CACGTTAATGGCCCTCATGATTGGCTTTTACGCTCTTTCGCCATTAAGTTCACGAAGT 465  
  
 QY 631 CAAGCTGATCTAGATGCGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 690  
 DB CAAGCTGATCTAGATGCGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 525  
  
 QY 691 GGCCTGCCCTTATCTTCTGAACTGACACCTCAACTGCCCCCATCGAAGCCCTTTCCTGCCCC 750  
 DB GGCCTGCCCTTATCTTCTGAACTGACACCTCAACTGCCCCCATCGAAGCCCTTTCCTGCCCC 585  
  
 QY 751 TGAGGAAGTCAATTAACAAGAAAGATCAAGCTGAGTGGCTGGGCTTGGCCCTTGACACCAAGGT 810  
 DB TGAGGAAGTCAATTAACAAGAAAGATCAAGCTGAGTGGCTGGGCTTGGCCCTTGACACCAAGGT 645  
  
 QY 811 GACAGGTGACCTTCTTACACCATGTGACCAACCATGGGCCAGAGGCTCAGCCAGAGGC 870  
 DB GACAGGTGACCTTCTTACACCATGTGACCAACCATGGGCCAGAGGCTCAGCCAGAGGC 705  
  
 QY 871 CCCCAGCCTGGAGGACGCTTCCGATGCTTCACTGTCACCCAGGATGTTTCGGGGCACCTC 930  
 DB CCCCAGCCTGGAGGACGCTTCCGATGCTTCACTGTCACCCAGGATGTTTCGGGGCACCTC 765  
  
 QY 931 AGAAAACTTCTGAGAGGCTCTGTCCCTTACGCAAGAGCTCTGTCCCCCACTTTCCT 990  
 DB AGAAAACTTCTGAGAGGCTCTGTCCCTTACGCAAGAGCTCTGTCCCCCACTTTCCT 825  
  
 QY 991 GTGGAGCCTCTTCAACCATGGGCGATGACCCAGCTGCGGATCATCTTCTACATGCTGCTGT 1050  
 DB GTGGAGCCTCTTCAACCATGGGCGATGACCCAGCTGCGGATCATCTTCTACATGCTGCTGT 885  
  
 QY 1051 GAACAGATGCTGGAGTACCTTGTGACTGGTGCCAGAGGATGAGACAAATGAACAGCA 1110  
 DB GAACAGATGCTGGAGTACCTTGTGACTGGTGCCAGAGGATGAGACAAATGAACAGCA 945  
  
 QY 1111 ACAAAGGTGGCAGAGACAGTTGGGTTTACTCTCCGTTTCGCGGCGCATGACAGTGT 1170  
 DB ACAAAGGTGGCAGAGACAGTTGGGTTTACTCTCCGTTTCGCGGCGCATGACAGTGT 1005  
  
 QY 1171 GTGCGCTTCTACCTGCCCTCATTTGGCTTACATGAGCTGCGGATCAAGAGCTGCGT 1230  
 DB GTGCGCTTCTACCTGCCCTCATTTGGCTTACATGAGCTGCGGATCAAGAGCTGCGT 1065  
  
 QY 1231 GGACGCCCCAACTCAGGGGCACTGTCTCGGAGATGCCAGGAGCGGGTTGCTACCAATC 1290  
 DB GGACGCCCCAACTCAGGGGCACTGTCTCGGAGATGCCAGGAGCGGGTTGCTACCAATC 1125  
  
 QY 1291 CATCAGACCACTGCTCAAGATCCAAAGCTCAACAAATGCAATGCAATGCAATGCAATGCAAT 1350  
 DB CATCAGACCACTGCTCAAGATCCAAAGCTCAACAAATGCAATGCAATGCAATGCAATGCAAT 1185  
  
 QY 1351 GACCAACCTGCTGCTTGGGTTTGGCATCACTGCTCTCATCAACAACTTACACCTCCA 1410  
 DB GACCAACCTGCTGCTTGGGTTTGGCATCACTGCTCTCATCAACAACTTACACCTCCA 1245  
  
 QY 1411 GTTTGTGACTTGTCTGTCACACCAATGTTGAGGTTTCTTCCACTCAGACCTTGGGAG 1470  
 DB GTTTGTGACTTGTCTGTCACACCAATGTTGAGGTTTCTTCCACTCAGACCTTGGGAG 1305  
  
 QY 1471 TCTATATGTCAGTGTTCCTTCCCATCCAAACCACTTTTGGAGCGCTGACAGGCTGCAAGTCCCT 1530  
 DB TCTATATGTCAGTGTTCCTTCCCATCCAAACCACTTTTGGAGCGCTGACAGGCTGCAAGTCCCT 1365  
  
 QY 1531 CATCAGTGTGTTGGCTTGGCTTGGCATCACTTTCATGGCCATGGTGGGACCCCT 1590  
 DB CATCAGTGTGTTGGCTTGGCTTGGCATCACTTTCATGGCCATGGTGGGACCCCT 1425  
  
 QY 1591 GAAAGGAGAGCCCTTCTGGGTGAATCTGGGCTCTCTGCTATTCTACTCTCTGGGATTCCT 1650











Query Match	29.4%;	Score 684.2;	DB 10;	Length 143899;
Best Local Similarity	97.9%;	Pred. No. 9.6e-201;		
Matches 704;	Conservative	0;	Mismatches 13;	Indels 2; Gaps 1;

  

Qy	1602	CCCTTCGGGTGAATCTGGGCCTCCTCTAFTCTCACTCTGGGATTCCTGTGGCTTCCT	1661
Db	38731	CCCTAAAGGTGAATCTGGGCCTCCTGCTATTCCTACTCTGGGATTCCTGTGGCTTCCT	38672
Qy	1662	ACCTCTTCTATTAACCGTCCCGCTCAGCAGAGGTACGCCGCCAATGGGATGGGCCAC	1721
Db	38671	ACCTCTTCTATTAACCGTCCCGCTCAGCAGAGGTACGCCGCCAATGGGATGGGCCAC	38612
Qy	1722	TGAAGTGCTTACGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGAGACCTGGAT	1781
Db	38611	TGAAGTGCTTACGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGAGACCTGGAT	38552
Qy	1782	GACGGCAATCAAGGCTTGAGCAACAAAGGAGTGCCCCATATGGCTTTTCTACCTGTA	1841
Db	38551	GACGGCAATCAAGGCTTGAGCAACAAAGGAGTGCCCCATATGGCTTTTCTACCTGTA	38492
Qy	1842	ACATGCACATAGAGCCATGGCCGTAGATTATATAATACCAAGAGAGTTCTATTTTTGTA	1901
Db	38491	ACATGCACATAGAGCCATGGCCGTAGATTATATAATACCAAGAGAGTTCTATTTTTGTA	38432
Qy	1902	AAGACTGCAAAAAGGAGGAAAAAAACCTTCAAAAAAGCCCCCTTAAGTCAACGCTCCATT	1961
Db	38431	AAGACTGCAAAAAGGAGGAAAAAAACCTTCAAAAAAGCCCCCTTAAGTCAACGCTCCATT	38372
Qy	1962	GACTGGAAGACAGTCCCTATCTCTAGAGGGGTGAGCTTTCTCTCCTCTGGGTGGAGAG	2021
Db	38371	GACTGGAAGACAGTCCCTATCTCTAGAGGGGTGAGCTTTCTCTCCTCTGGGTGGAGAG	38312
Qy	2022	ACCAGGTGCCTCTTATCTCTTCTAGCGGTCTGCTCCTCTGTACTCTTGGGGGATCG	2081
Db	38311	ACCAGGTGCCTCTTATCTCTTCTAGCGGTCTGCTCCTCTGTACTCTTGGGGGATCG	38252
Qy	2082	GCAACAGGCTACCCCTGAGTCCCATGTGCCATGAGTGTGCACACATGCAATGTGTCT	2141
Db	38251	GCAACAGGCTACCCCTGAGTCCCATGTGCCATGAGTGTGCACACATGCAATGTGTCT	38194
Qy	2142	GTGTATGTGTGAATGTGAGAAAAACACAGCCCTCTCTTTCAGAAAGGAAAGGGCCCTGAGGT	2201

Query Match	18.2%	Score 423.6;	DB 9;	Length 464;
Best Local Similarity	97.4%;	Pred. No. 6.1e-121;		
Matches 452;	Conservative 0;	Mismatches 9;	Indels 3;	Gaps 2;
Qy	1863	CGTAGATTTTATAAATACCAAGAGAAGTTCATTTTTGTTAAAGACTGCAAAAGAGGAGAAA	1922	
Db	464	CGTAGATTTTATAAATACCAAGAGAAGTTCATTTTTGTTAAAGACTGCAAAAGAGGAGAAA	405	
Qy	1923	AAAAACCTTCAAAAACGCCCCCTTAAGTCAACGCTCCATTGACTGAAGACAGTCCCTATCC	1982	
Db	404	AAAAACCTTCAAAAACGCCCCCTTAAGTCAACGCTCCATTGACTGAAGACAGTCCCTATCC	345	
Qy	1983	TAGAGGGTTGAGCTTTCTTCCTCTTGGGGTGGAGAGACAGGGTGCTCTTATCTCC	2042	
Db	344	TAGAGGGTTGAGCTTTCTTCCTCTTGGGGTGGAGAGACAGGGTGCTCTTATCTCC	285	
Qy	2043	TTCTAGGGGTCGCCCTCTGGTACCTCTTGGGGGATCGCAACACAGGCTACCCCTGAGG	2102	
Db	284	TTCTAGGGGTCGCCCTCTGGTACCTCTTGGGGGATCGCAACACAGGCTACCCCTGAGG	235	
Qy	2103	TCCCATGTGCCATGAGTGTGCACAAATGTCATGTGTCTGTGTATGTGTGAATGTGAGAA	2162	
Db	224	TCCCATGTGCCATGAGTGTGCACA--CATGCATGTGTCTGTGTATGTGTGAATGTGAGAG	167	
Qy	2163	AAACACAGCCCTCTCTTTCAGAGAAAGGGCCTGAGGTGCCAGCTGTGTCTGGGTTAG	2222	
Db	166	AGACACAGCCCTCTCTTTCAGAGAAAGGGCCTGAGGTGCCAGCTGTGTCTGGGTTAG	107	
Qy	2223	GGGTTGGGGTGGCCCCCTTCAGGGGCCAGGAGGAGGTTCCCTCTCTGGTGCTGCTGC	2282	
Db	106	GGGTTGGGGTGGCG-CCCTTCAGAGGCCAGGAGGGCAGGTTCCCTCTCTGGTGCTGCTGC	48	
Qy	2283	TTGCAAGCTCTTAGAGAAATAAAAAGGAAAGTGCAGAAAAA	2326	
Db	47	TTGCAAGCTCTTAGAGAAATAAAAAGGAAAGTGCAGAAAAA	4	

RESULT 13

US-09-880-1107-539/c



```
Db      94  GAGCATGAGACAAATGAACAGCAACAAAGGTGGCAGAGACAGTTGGGTTTACTCCTCC 153
Qy      1148 GTCTTCGGGGCCATGCAGCTGTGTGCTTCTCACCTGCCCTCATTTGGCTACATCATG 1207
Db      154  GTCTTCGGGGCCATGCAGCTGTGTGCTTCTCACCTGCCCTCATTTGGCTACATCATG 213
Qy      1208 GACTGGCGGATCAAGGACTGCGTGAGACGGCCCAACTCAGGGCACTGTCTCGGAGATGCC 1267
Db      214  GACTGGCGGATCAAGGACTGCGTGAGACGGCCCAACTCAGGGCACTGTCTCGGAGATGCC 273
Qy      1268 AGGACGGGTTGCTTACCAATCCATCAGACCAAGCTACTGCAAGATCCAAAAGCTCACC 1327
Db      274  AGGACGGGTTGCTTACCAATCCATCAGACCAAGCTACTGCAAGATCCAAAAGCTCACC 333
Qy      1328 AATGCCATCAGTGCCTTACCCCTGACCAACCTGCTGCTGTGTGGGTTTGGCATCACCTGT 1387
Db      334  AATGCCATCAGTGCCTTACCCCTGACCAACCTGCTGCTGTGTGGGTTTGGCATCACCTGT 393
Qy      1388 CTCATCAACAATTACACCTCC 1409
Db      394  CTCATCAACAATTACACCTCC 415
```

Search completed: April 10, 2004, 20:21:38  
Job time : 924.185 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:16:47 ; Search time 7925.34 Seconds  
(without alignments)  
8764.220 Million cell updates/sec

Title: US-09-743-825-1  
Perfect score: 2326  
Sequence: 1 ccggcgctggagggggcga.....agggaagtggagaaaaaaa 2326

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_estc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_nam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_pbg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1311.6	56.4	2490	11 AK031215	AK031215 Mus muscu
2	1282	55.1	2129	11 AK011417	AK011417 Mus muscu
C 3	897.8	38.6	982	9 AL580746	AL580746 AL580746
C 4	829.2	35.6	1201	9 AL581753	AL581753 AL581753

5	804	34.6	977	12	BM451394	BM451394
6	769.8	33.1	902	13	BUI56542	BUI56542
7	768.4	33.0	1058	12	BI758694	BI758694
8	760.2	32.7	830	12	BG468720	BG468720
9	750.8	32.3	785	29	AY417252	AY417252
10	750.6	32.3	884	13	BUI82024	BUI82024
11	747	32.1	839	10	BE396312	BE396312
C 13	742.8	31.9	785	29	AY417253	AY417253
14	728.6	31.3	933	13	BX392039	BX392039
15	697.8	30.0	726	12	BG821730	BG821730
16	690.4	29.7	837	12	BG823510	BG823510
17	681	29.3	972	10	BG497880	BG497880
18	677	29.1	677	13	BE250078	BE250078
19	665.2	28.6	1026	12	BX103969	BX103969
20	661.4	28.4	866	12	BG328396	BG328396
21	645.2	27.7	923	10	BG823832	BG823832
C 22	638.6	27.5	726	10	BE293435	BE293435
23	633.4	27.2	814	10	BE872354	BE872354
24	620.8	26.7	763	12	BF311736	BF311736
25	620.2	26.7	823	10	BI827416	BI827416
26	618	26.6	1201	9	BG032562	BG032562
27	612.6	26.3	1201	13	AL560277	AL560277
C 28	605.6	26.0	665	12	BX447962	BX447962
29	596.2	25.6	882	10	BQ006067	BQ006067
30	593	25.5	1092	12	BE396979	BE396979
31	589.4	25.3	702	10	BG328388	BG328388
C 32	586.6	25.2	656	12	BG396500	BG396500
33	582.6	25.2	815	10	BQ005970	BQ005970
34	582.2	25.0	679	10	BF664173	BF664173
35	578.4	24.9	764	10	BE868257	BE868257
C 36	573.2	24.6	1171	13	BE560932	BE560932
C 37	562.8	24.2	622	10	BX447961	BX447961
C 38	561.4	24.1	613	12	AW967482	AW967482
39	559.6	24.1	617	10	BM677811	BM677811
40	558.2	24.0	941	13	AW566866	AW566866
41	558.2	24.0	930	13	BUI51063	BUI51063
42	551.2	23.7	941	13	BQ889168	BQ889168
43	546.4	23.5	586	12	BQ715545	BQ715545
44	543.6	23.4	534	9	BM725383	BM725383
45	539.6	23.2	762	29	AV688214	AV688214
					AY417254	AY417254

## ALIGNMENTS

RESULT 1  
AK031215.  
LOCUS

DEFINITION

AK031215 2490 bp mRNA linear HTC 18-SEP-2003  
Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length  
enriched library, clone:5930430D04 product:prostate cancer  
overexpressed gene 1, full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AK031215.1 GI:26327122  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED

1  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999).  
99279253  
10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
PUBMED  
11042159



Db 1059 CATCCCGGATATCCCTGGTACCTCAGAGAGACTCCTGAAAGTCTGTCCCTTTTCGCA 1118  
Qy 966 AGAGCCTCTGCTCCCACTTTCTCTGAGGCTCTCACCATGGCAGCCAGCTGC 1025  
Db 1119 AGAGCCTCTGCTCCCACTTTCTCTGAGGCTCTCACCATGGCAGCCAGCTTC 1178  
Qy 1026 GGATCATCTTTACATGGCTGTGTGAACAAGATGCTGGAGTACCTTTGTGACTGGTGGCC 1085  
Db 1179 GGGTCATCTTTATATGGGTGCTATGAACAAGATCTGGAGTTCAATTGTGACTGGTGGCA 1238  
Qy 1086 AGAGCATGAGCAAAATGAACAGCAACAAAGGTGGCAGAGACAGATGGGTTCATCTCT 1145  
Db 1239 AGGAACGTGAGCAAAATGAGCAGACAGAAGGTGGAGGAGACAGTTGAGTTCTACTCTT 1298  
Qy 1146 CCGTCTTTCCGGGCCATGACAGCTGTGTGCTCTTCTACCTGCCCCCTCATTTGGCTACATCA 1205  
Db 1299 CCATCTTTGGAGTCATGACAGCTGTGTGCTCTTCTACCTGCCCCCTCATTTGGCTACATCA 1358  
Qy 1206 TGGACTGGCGGATCAAGGACTGCTGGAGCGCCCACTCAGGCACTGT----- 1254  
Db 1359 TGGACTGGCGCATCAAGGACTGTGTGATGCTCCACGGAGGGCACCTGAATGAGATG 1418  
Qy 1255 -----CTCGGAGATGCCAGGACGGGTTGCTACCAATCCATCAGACCAAGCTACTGCA 1310  
Db 1419 CTTCTCTTGGAGATGCCAGATGGGCTAGCACCAAGTTCACTAGACCAAGCTACTGCA 1478  
Qy 1311 AGATCCAAAGCTCCCAATGCCATCAGTCAGTCCCTTACCTGACCACTGCTGTGTGG 1370  
Db 1479 AGGTACAAAGCTCACCATGGCATCAATGCCCTTACCTGACCACTGCTGTGTGG 1538  
Qy 1371 GTTTTGGCATCACTGCTCATCAACAACTTACACCTCAGTTTGTGACTTTGTCTGCTG 1430  
Db 1539 GTTTGGCATCGCTGCTCATCAAGAACTTACACCTGAGTTGTGGCTTTGTCTGCTG 1598  
Qy 1431 ACACCATGTTGAGGTTTCTTCCATCAGCCTGTGGAGTCTCTATGTGAGTGTTC 1490  
Db 1599 ATACCATAGTTGCGGTTTCTTCCATCAGCCTGTGGAGTCTCTACGCTGTGTGTTC 1658  
Qy 1491 CATCCAACTTTGGAGCTGACAGGCTGACAGGCTGAGTCCCTCATCAGTGTGTTCGCT 1550  
Db 1659 CGTCCAACTATTTTGGGACACTGACAGGTTCTCAGTCTCTCATCAGTGTGTTCGCT 1718  
Qy 1551 TGCTTCAGCAGCCACTTTTTCATGGCATGTGTGGACCCCTGAAAGAGAGCCCTTCGGG 1610  
Db 1719 TGCTGCAACAGTACTCTTCTATGGCCATGTGGGACCCCTGCATGAGATCCCTTCGG 1778  
Qy 1611 TGAATCTGGCCTCTGCTATTCTCTACTCTGGGATTCCTGTGTGCTTCTACCTCTTCT 1670  
Db 1779 TGAACCTGGGCTCTTCTCTCTGTTCTCTGGAATTTCTCTACCTTCTTCTACTCT 1838  
Qy 1671 ATTACGTCGGGCTCCAGCAGGATAGCCGCCAATGGGATGGGCCCTCAGTGAAGTGC 1730  
Db 1839 ACTACCGGCTCGCCCTGCAAGAGAGATATGCCACCAATTTGGTAGACCCACAGAGGTC 1898  
Qy 1731 TTAGCGGCTCTGAGGTGACCGATAGACTTCTCAGACCAAGGACCTGGATGACAGGCAA 1790  
Db 1899 TCATATCTTGAAGTGGCTACATAGACTCTCTGAGGCCAAGAGACTTGGAGNACGCGAG 1958  
Qy 1791 TCAAGCCTGAGCAACCAAAAGAGTGCCCAATATGGCTTTTCTACCTGTAAATGACACA 1850  
Db 1959 TCAAGCCTGATAAACCAGGAGGAAATGGGCTGT---GGCTTTCTACCTGCATCGTGTCA 2015  
Qy 1851 TAGAGCCATG-CGCGTAGATTATTAATACCAAGAGAGTCTATTTTGTAAAGACTGC 1909  
Db 2016 TAGAGCCGGTCTGTGGATTTATAAATACTAAG---AGTTCTATTTTGTAGGGGCTG 2072  
Qy 1910 AAAAGGAGGAAAAA-----CCTTCAAAACGCC 1942  
Db 2073 CAAAAAGGAAACAAACAGAACAAACCCCAACCAAAACAGCCCTTAAAAAATCCC 2132  
Qy 1943 CCTAAGTCAACGCTCCTGACTGAAGACAGTCCCTTATCTTACAGGGGTTGAGCTTCTT 2002

Db 2133 CCCAACCAAGCTATCCATCACTGAAGACAGTGCCTGTTCCAGTGGTATTGGGCCATCGT 2192  
Qy 2003 CCTCCTTGGGTTGGAGGAGACAGGCTGCTCTTATCTCTTCTAGCGGCTGCTCCCTG 2062  
Db 2193 CTTTGAAGAAAGTGTGGGC-----TGCCCTTTTCCACACTTGGCCTAG 2239  
Qy 2063 GTACCTCTTTGGGGGATCGGCAACAGGCTACCCCTGAGGTCCCATGTGCCATGAGTGTG 2122  
Db 2240 GAGCCTCTGTGGAGTTGCTTCTGAAAACTCATGGATCAGTGAACAGACACACATGA 2299  
Qy 2122 CACAACATGCAATGTCTGTGTATGTGTGAATGTGAGAAACACAGCCCTCTTTCAG 2182  
Db 2300 TTCCAGCCGCAAGTGTGCACATATATTTACGCACATGCATATTCATCTCTTTCCTG 2359  
Qy 2183 AAGGAAAGGGGCTCAGGTGCGAGTGTCTCTGCTGTAGGGTGTGGGGTGGGCCCTT 2242  
Db 2360 CAGGAGA-AGGACTGAGGGCTGGCTGTATCCAGATGGGTGTGGGG----- 2407  
Qy 2243 CCAGGGCAGGAGGAGGCTTCCCTCTCTGTGTGCTGCTTGTCAAGTCTTAGAGGAAT 2302  
Db 2408 -----CTGGAGGGCGGTTCCCTCCCTGATGCTGTTCTTACAGTCTTAGAGGAAT 2460  
Qy 2303 AAAAAGGGAATGAGA 2318  
Db 2461 AAAAAGGGAATGAAA 2476  
RESULT 2  
AK011417 2129 bp mRNA linear HTC 20-SBP-2003  
LOCUS Mus musculus 10 days embryo whole body cDNA, RIKEN full-length  
DEFINITION enriched library, clone:2610016F07 product:prostate cancer  
overexpressed gene 1, full insert sequence.  
ACCESSION AK011417.1 GI:12847526  
VERSION HTC; CAP trapper.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
REFERENCE 1 Carninci, P. and Hayashizaki, Y.  
AUTHORS High-efficiency full-length cDNA cloning  
TITLE Meth. Enzymol. 303, 19-44 (1999)  
JOURNAL 99279253  
MEDLINE 10349636  
PUBMED  
REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
AUTHORS Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes  
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)  
PUBMED 20499374  
REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
AUTHORS Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Izawa, M., Ohara, E., Watanabe, M.,  
Yoneda, Y., Ishikawa, T., Ogawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multipillar sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11042159  
REFERENCE 4 The RIKEN Genome Exploration Research Group Phase II Team and the  
AUTHORS PANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
PUBMED 11076861  
REFERENCE 5





1086 AGGAGCATGAGCAAAATGAACAGCAAAAGGTGGCAGAGACAGTGTGGGTTCTACTCT 1145  
 1295 AGGAAGTGGAGCAAAATGAGCAGAGACAGAGGTGGAGGAGAGTGTGTTCTACTCT 1354  
 1146 CGCTCTTGGGGCCATGAGCTGTGTGCTTCTACCTGCCCCCTCATTTGGCTACATCA 1205  
 1355 CCATCTTGGAGTCAATGAGCTGTGTGCTTCTACCTGCCCCCTCATTTGGCTACATCA 1414  
 1206 TGGACTGGCGGATCAAGGACTGGTGGAGCGCCCAACTCAGGCA----- 1250  
 1415 TGGACTGGCGCATCAAGGACTGTGTGATGCTCAACGAGGGCACCTGAATGAGAAATG 1474  
 1251 CTGTCTTCGAGATGCCAGGAGCGGGTTGCTACCAAAATCCATCAGACCAAGCTACTGCA 1310  
 1475 CTTCTTTGGAGATGCCAGATGGGCTAGCACCAGTTCATAGACCAAGCTACCGCA 1534  
 1311 AGATCAAAAGCTCAACATGCAATGCAATGCTTCACTTCACTTCACTTCACTTCACTT 1370  
 1535 AGGTACAAAGCTCACCAGTGCATCAATGCTTCACTTCACTTCACTTCACTTCACTT 1594  
 1371 GTTTTGGCATCACTGTCTCATCAACAACTTACACCTCCAGTTTGTGACTTTGTCTCTG 1430  
 1595 GTTTTGGCATCGCTCGCTCATCAAGAACTTACACCTGAGTTGTGGCTTTGTCTCTG 1654  
 1431 ACACCAATGTTGAGGTTCTTCCACTCAGCTGTGGAGTCTCTATGCTGCAAGTCTCC 1490  
 1655 ATACCATAGTTCGCGTTCTTCCACTCAGCTGTGGAGTCTCTACGCTGCTGTCTTC 1714  
 1491 CATCAACCACTTTGGAGCGCTGACAGGCTGCGAGTCCCTCATCATGCTGTGTGCTCT 1550  
 1715 CGTCCCAATCATTTTGGGACACTGACAGGCTTTCAGTCTCTCATCATGCTGCGTGT 1774  
 1551 TGCTTCAGAGCCACTTTTCATGCGCATGTTGGGACCCCTGAAAGAGAGCCCTCTGG 1610  
 1775 TGCTGCAACAGCTACTCTTATGCGCATGTTGGGACCCCTGCAATGAGAGTCCCTCTGG 1834  
 1611 TGAATCTGGGCTCTCTGCTTCTCTACTCTGCGATTTCTGTTCTCTCTCTCTCTCT 1670  
 1835 TGAACCTGGGCT 1894  
 1671 ATTACGCTGCGGCTTCAGCAGGAGTACGCGCCCAATGGGATGGGCCCACTGAAGTGC 1730  
 1895 ACTACCGGCTCTGCGCTGCAAGAGAGTATGCCACCAATTTGGTAGACCCACAGAGTGC 1954  
 1731 TTAGCGGCTCTGAGTGCGCATAGCTTCTCAGACCAAGGACCTGATGACAGGCA 1790  
 1955 TCAATCTTGAAGGTGGCTACATAGACTCTGAGGCCAAGAGACTTGGAGGACAGCAG 2014  
 1791 TCAAGGCTGAGCAACAAAGAGTGGCCCATATGCTTTTCTACCTGTAAACATGCACA 1850  
 2015 TCAAGGCTGATAAACCGAAGGAATGGCTGT---GGCTTTCTACCTGCATCGTGTCA 2071  
 1851 TAGAGCCATG-GCCGTAGATTATTAATACCAAGAGAGTCTTATTTTGTAAAGACT 1907  
 2072 TAGAGCCGGGTTCTGTGGATTATTAATAACTAAG---AGTTCTATTTTGTAGGACT 2126

RESULT 3  
 AL580746/c 982 bp mRNA linear EST 01-JUN-2003  
 LOCUS  
 DEFINITION AL580746 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED  
 Homo sapiens cDNA clone CS0DJ010YL17 3-PRIME, mRNA sequence.  
 ACCESSION AL580746  
 VERSION  
 KEYWORDS EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 982)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)

COMMENT

On Feb 16, 2001 this sequence version replaced gi:12947068.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 6149.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DJ010CF09NP1&cluster=6149.r. Contact :  
 Feng Liang Email: fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Paraday Avenue Genoscope sequence ID : CS0DJ010CF09NP1.

FEATURES

Source

1. 982  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DJ010YL17"  
 /cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
 /cell\_line="JURKAT"  
 /clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT  
 10-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 38.6%; Score 897.8; DB 9; Length 982;  
 Best Local Similarity 96.1%; Pred. No. 3.5e-192;  
 Matches 910; Conservative 15; Mismatches 20; Indels 2; Gaps 1;  
 QY 1314 TCACAAAGCTACCAATGCCATCAGTGCCTTACACCTGACCAACCTGCTGTGGGTT 1373  
 DB 982 TCACAAAGCTACCAATGCCATCAGTGCCTTACACCTGACCAACCTGCTGTGGGTT 923  
 QY 1374 TTGGGATCACTGTCTCATCAACAACTTACACCTCAGTTTGTGACCTTTGTCTGCACA 1433  
 DB 922 TTGGGATCACTGTCTCATCAACAACTTACACCTCAGTTTGTGACCTTTGTCTGCACA 863  
 QY 1434 CAAATTTGAGGTTTCTTCCACTCAGCTGTGGAGTCTCTATGCTGCAAGTGTCCCAT 1493  
 DB 862 CAAATTTGAGGTTTCTTCCACTCAGCTGTGGAGTCTCTATGCTGCAAGTGTCCCAT 803  
 QY 1494 CCAACCACTTTGGGAGCGTACAGGCTGCAAGCTCCTCATCAGTGTGTGGCTTGC 1553  
 DB 802 CCAACCACTTTGGGAGCGTACAGGCTGCAAGCTCCTCATCAGTGTGTGGCTTGC 743  
 QY 1554 TTCAAGGCTGAGCAACAAAGAGTGGCCCATATGCTTTTCTACCTGTAAAGACT 1613  
 DB 742 TTCAAGGCTGAGCAACAAAGAGTGGCCCATATGCTTTTCTACCTGTAAAGACT 683  
 QY 1614 ATCTGGGCTCTCTGCTATTTCTCACTCCTGGGATTCCTGTGCTTCTCTACCTTCTATT 1673  
 DB 682 ATCTGGGCTCTCTGCTATTTCTCACTCCTGGGATTCCTGTGCTTCTCTACCTTCTATT 623  
 QY 1674 ACCGTGCGGCTCCAGCAGGAGTACGCGCCCAATGGGATGGCCCACTGAAGTGTCTTA 1733  
 DB 622 ACCGTGCGGCTCCAGCAGGAGTACGCGCCCAATGGGATGGCCCACTGAAGTGTCTTA 563  
 QY 1734 GGGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGAGCTTGATGACAGGCAATCA 1793  
 DB 562 GGGGCTCTGAGGTGACCGCATAGACTTCTCAGACCAAGGAGCTTGATGACAGGCAATCA 503  
 QY 1794 AGGCTGAGCAACCAAAAGGAGTGGCCCATATGCTTTTCTACCTGTAAAGACTGCAATAG 1853  
 DB 502 AGGCTGAGCAACCAAAAGGAGTGGCCCATATGCTTTTCTACCTGTAAAGACTGCAATAG 443  
 QY 1854 AGCCATGGCCGTAGATTTTATAATACCAAGAGAGTCTTATTTTGTAAAGACTGCAAAA 1913  
 DB 442 AGCCATGGCCGTAGATTTTATAATACCAAGAGAGTCTTATTTTGTAAAGACTGCAAAA 383



```

DEFINITION AGENCOURT_6390890 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493944
5', mRNA sequence.
ACCESSION BM451394
VERSION BM451394.1 GI:18500434
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 977)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
COMMENT NIH-GMC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12118 row: h column: 09
High quality sequence stop: 614.
Location/Qualifiers
1..977
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5493944"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

FEATURES
source
Query Match 34.6%; Score 804; DB 12; Length 977;
Best Local Similarity 98.3%; Pred. No. 66-171;
Matches 855; Conservative 0; Mismatches 10; Indels 5; Gaps 4;

QY 189 TCATTCGAAAGACGAGGCTTCTATTCCAGCAGCTGCCAGCTGAGAGCAGCACCACCA 248
DB 1 TCATTCGAAAGACGAGGCTTCTATTCCAGCAGCTGCCAGCTGAGAGCAGCACCACCA 60

QY 249 CCACCCAGGATGACGACGAGTGCCAGGCTGTGACCCAGCAGGACGAGATGCTCAACC 308
DB 61 CCACCCAGGATGACGACGAGTGCCAGGCTGTGACCCAGCAGGACGAGATGCTCAACC 120

QY 309 TGGGCTTACCATGGTCTTCTGCTGCTCAGCGCCACACCTGCCACTGGGGATCCTCA 368
DB 121 TGGGCTTACCATGGTCTTCTGCTGCTCAGCGCCACACCTGCCACTGGGGATCCTCA 180

QY 369 TGGACCGCTTTGGCCCCCAGCCGCTGGCTGGCAGTGTCTCTCACTGGGCTCT 428
DB 181 TGGACCGCTTTGGCCCCCAGCCGCTGGCTGGCAGTGTCTCTCACTGGGCTCT 240

QY 429 GCACCTCATATGGCCCTGGGCTCCCGGACGTGAAGCTGTCTCCGTTGATATTCCTGG 488
DB 241 GCACCTCATATGGCCCTGGGCTCCCGGACGTGAAGCTGTCTCCGTTGATATTCCTGG 300

QY 489 CGCTGTCCCTGATGGCTTTGGTGATCTGCTTAAGCTTCACTTCACTCAGCTGCCCA 548
DB 301 CGCTGTCCCTGATGGCTTTGGTGATCTGCTTAAGCTTCACTTCACTCAGCTGCCCA 360

QY 549 ACATGTTTGGGAACCTGGCTCCACGCTTAATGGCCCTCATGATTGGCTCTTACGCTCT 608
DB 361 ACATGTTTGGGAACCTGGCTCCACGCTTAATGGCCCTCATGATTGGCTCTTACGCTCT 420

QY 609 CTGCCATTACGTTCCAGGAATCAAGCTGATCTACGATCCCGGTGTGGCTTCTGGTCA 668

```

```

DB 421 CTGCCATTACGTTCCAGGAATCAAGCTGATCTACGATCCCGGTGTGGCTTCTGGTCA 480
QY 669 TCATGTTTACCTGGTCTGGCCCTGCTTATCTTTCTGAACCTGACCCCTCAACTGSC 728
DB 481 TCATGTTTACCTGGTCTGGCCCTGCTTATCTTTCTGAACCTGACCCCTCAACTGSC 540
QY 729 CCATCGAAGCTTTCCTGCCCTCGAGGAAGTCAATTTACACGAAGAGATCAAGCTGAGTG 788
DB 541 CCATCGAAGCTTTCCTGCCCTCGAGGAAGTCAATTTACACGAAGAGATCAAGCTGAGTG 600
QY 789 GCGTGGCCCTGACACCAAGGTGACAGGTGACCTTTTACACCCATGTGACACCATGG 848
DB 601 GCGTGGCCCTGACACCAAGGTGACAGGTGACCTTTTACACCCATGTGACACCATGG 660
QY 849 GCACGAGGCTCAGCAGAGAGGCCCCAGCCTCGAGGACGGTTCGGATGCTTCATGCTCAC 908
DB 661 GCACGAGGCTCAGCAGAGAGGCCCCAGCCTCGAGGACGGTTCGGATGCTTCATGCTCAC 720
QY 909 CCCAGGATGTTCCGGGACCTTCAGAAAACCTTCTCGAGAGGTCTGTCCCC-TTACGCAAG 967
DB 721 CCCAGGATGTTCCGGGACCTTCAGAAAACCTTCTCGAGAGGTCTGTCCCC-TTACGCAAG 780
QY 968 AGCCTCTG-CTCCCCACCTTCTCTGTGAGGCTCTCTCACCATGGG--CATGACCCAGCTG 1024
DB 781 AGCCTCTGCTCCCCACCTTCTCTGTGAGGCTCTCTCACCATGGGCGCATGACCAAGCTG 840
QY 1025 CGGATCATCTT-CTACATGGCTGTGTGAA 1053
DB 841 CGGATCATCTTCTCTAAATGGCTGTGTGAA 870

RESULT 6
BUI56542
LOCUS AGENCOURT_7859106 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6168698
DEFINITION 5', mRNA sequence.
ACCESSION BUI56542
VERSION BUI56542.1 GI:22670074
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 902)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13532 row: k column: 03
High quality sequence stop: 575.
Location/Qualifiers
1..902
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6168698"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

FEATURES
source
ORIGIN

```



Db 523 ATTCTGGCGCTGTCCCTGAATGGCTTTGGTGGCACTCTCCCTAAAGCTTCACTCTCACTCAC 582

QY 541 GCTGCCCAACATGTTTGGGAACCTGGCTCCACGTTAAATGGCCCTCATGATTGGCTCTTA 600

Db 583 GCTGACCAACATGTTTGGGAACCTGGCTCCACGTTAAATGGCCCTCATGATTGGCTCTTA 642

QY 601 CGCTCTTCTGCAATTAAGTTCACAGGAATCAAGCTGATCTACGATGCCGGTGTGGCCTT 660

Db 643 CGCTCTTCTGCAATTAAGTTCACAGGAATCAAGCTGATCTACGATGCCGGTGTGGCCTT 702

QY 661 CTGGGTGATCATGTTTCACTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 720

Db 703 CTGGGTGATCATGTTTCACTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 761

QY 721 CAACTGGCCCATCGAA-GCCTTTCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 778

Db 762 CAACTGGCCCATCGAA-GCCTTTCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 821

QY 779 AAGCTGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 812

Db 822 AAGCTGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 855

RESULT 8

BG468720 830 bp mRNA linear EST 21-MAR-2001

LOCUS 602510136F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4644512 5',

DEFINITION mRNA sequence.

ACCESSION BG468720.1 GI:13400995

VERSION BG468720.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 830)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: [cgapbs-rc@mail.nih.gov](mailto:cgapbs-rc@mail.nih.gov) Tissue Procurement: ARCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: NIH Intramural Sequencing Center Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Plate: LLCM1416 row: g column: 09 High quality sequence stop: 825. Location/Qualifiers 1..830 /organism="Homo sapiens" /mol\_type="mRNA" /db\_xref="taxon:9606" /clone="IMAGE:4644512" /issue\_type="adenocarcinoma cell line" /lab\_host="DH10B (phage-resistant)" /clone\_lib="NIH\_MGC\_15" /note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"

ORIGIN

Query Match 32.7%; Score 760.2; DB 12; Length 830;

Best Local Similarity 99.4%; Pred. No. 4.7e-161;

Matches 784; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 CCGGGGCTGGAGGGGGCAAGCGGGTTCGAGGTGCAAGCCCTGGTGGCCCGAGCCCTGC 60

Db 44 CCGGGGCTGGAGGGGGCAAGCGGGTTCGAGGTGCAAGCCCTGGTGGCCCGAGCCCTGC 103

QY 61 GGAGCTCGGGGGCAGCATGGCCCCCAGCTGCAACAGGGCTACCGGAGGGCTGGTGGAT 120

Db 104 GGAGCTCGGGGGCAGCATGGCCCCCAGCTGCAACAGGGCTACCGGAGGGCTGGTGGAT 163

QY 121 GGCTGTCACGGCTGTGCTGGAGAACTCTTCTTCTGCTGTACTCTCTGGGCTGGGGCTC 180

Db 164 GGCTGTCACGGCTGTGCTGGAGAACTCTTCTTCTGCTGTACTCTCTGGGCTGGGGCTC 223

QY 181 CCTGTGATCATTTCTGAAGAACGAGGGCTTTATTCCAGCACGTGCCAGCTGAGAGCAG 240

Db 224 CCTGTGATCATTTCTGAAGAACGAGGGCTTTATTCCAGCACGTGCCAGCTGAGAGCAG 283

QY 241 CACCAACACCCAGGATGAGCGGAGGGTGGCAGGCTGTGACCAAGAGAGAGAT 300

Db 284 CACCAACACCCAGGATGAGCGGAGGGTGGCAGGCTGTGACCAAGAGAGAGAT 343

QY 301 GCTCAACCTGGGCTTTCACCATTTGGTTCTTCTGCTGTGCTCAGCGCACCACTGGCTGG 360

Db 344 GCTCAACCTGGGCTTTCACCATTTGGTTCTTCTGCTGTGCTCAGCGCACCACTGGCTGG 403

QY 361 GATCTCATGAGACCGCTTTGGCCCCCGACCCCTGGCGGCTGGTGGCAGTGCCTGCTTCA 420

Db 404 GATCTCATGAGACCGCTTTGGCCCCCGACCCCTGGCGGCTGGTGGCAGTGCCTGCTTCA 463

QY 421 TGGCTCTGACCCCTCATGGCCCCCTGGCCCTCCCGGACGTGGAAGCTCTGTCTCCGTTGAT 480

Db 464 TGGCTCTGACCCCTCATGGCCCCCTGGCCCTCCCGGACGTGGAAGCTCTGTCTCCGTTGAT 523

QY 481 ATTCTGGCGCTGTCCCTGAATGGCTTTGGTGGCACTCTGCCTAAAGCTTCACTCTCACTCAC 540

Db 524 ATTCTGGCGCTGTCCCTGAATGGCTTTGGTGGCACTCTGCCTAAAGCTTCACTCTCACTCAC 583

QY 541 GCTGCCCAACATGTTTGGGAACCTGGCTCCACGTTAAATGGCCCTCATGATTGGCTCTTA 600

Db 584 GCTGCCCAACATGTTTGGGAACCTGGCTCCACGTTAAATGGCCCTCATGATTGGCTCTTA 643

QY 601 CGCTCTTCTGCAATTAAGTTCACAGGAATCAAGCTGATCTACGATGCCGGTGTGGCCTT 660

Db 644 CGCTCTTCTGCAATTAAGTTCACAGGAATCAAGCTGATCTACGATGCCGGTGTGGCCTT 703

QY 661 CTGGGTGATCATGTTTCACTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 720

Db 704 CTGGGTGATCATGTTTCACTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 763

QY 721 CAACTGGCCCATCGAAAGCGCTTTCTGCTGGCTGGAGAGTCAATTACAGGAAGATCAA 780

Db 764 CAACTGGCCCATCGAAAGCGCTTTCTGCTGGCTGGAGAGTCAATTACAGGAAGATCAA 821

QY 781 GCTGAGTGG 789

Db 822 GCTGAGTGG 830

RESULT 9

AY417252 785 bp DNA linear GSS 17-DEC-2003

LOCUS Homo sapiens POVI gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

DEFINITION AY417252

ACCESSION AY417252.1 GI:39773212

VERSION GSS.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 785)

AUTHORS Clark,A.G., Ghanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,



Db 241 TCACTCCTGGATTCCTGTGCTTCTACCTCTCTTATTACGTCGCCGCTCCAGCAG 300

QY 1694 GAGTAGCGCCCAATGGGATGGGCCCACTGAAGGTGCTTAGCGGCTCTGAGGTGACCGCA 1753

Db 301 GAGTAGCGCCCAATGGGATGGGCCCACTGAAGGTGCTTAGCGGCTCTGAGGTGACCGCA 360

QY 1754 TAGACTTCTCAGACCAAGGACCTGTGATGACAGGCAATCAAGGCTTGAGCAACCAAAAGG 1813

Db 361 TAGACTTCTCAGACCAAGGACCTGTGATGACAGGCAATCAAGGCTTGAGCAACCAAAAGG 420

QY 1814 AGTGCCCATATGGCTTTCTACCTGTAACTGATGACATGACATGACGCTGCGGTAGATTAT 1873

Db 421 AGTGCCCATATGGCTTTCTACCTGTAACTGATGACATGACATGACGCTGCGGTAGATTAT 480

QY 1874 AAATACCAAGAGAGTCTTATTTTCTAAAGACTGCAAAAAGGAGGAAAAAACCCTTCA 1933

Db 481 AAATACCAAGAGAGTCTTATTTTCTAAAGACTGCAAAAAGGAGGAAAAAACCCTTCA 540

QY 1934 AAAACGCCCTTAAGTCAACGCTTCAATTGACTGAAGACAGTCCCTATCTAGAGGGTTG 1993

Db 541 AAAACGCCCTTAAGTCAACGCTTCAATTGACTGAAGACAGTCCCTATCTAGAGGGTTG 600

QY 1994 AGCTTCTTCTCCTTGGCTTGAGGACGACGAGGCTCTTATCTCTTCTAGCGGTC 2053

Db 601 AGCTTCTTCTCCTTGGCTTGAGGACGACGAGGCTCTTATCTCTTCTAGCGGTC 660

QY 2054 TGCTCTCTGTACTC- TTGGGGGATCGGCAACAGGCTACCCCTGAGGTCCAT- GTG 2111

Db 661 TGCTCTCTGTACTCTTTGNGGGGATCGGCAACAGGCTACCCCTGAGGTCCATGTG 720

QY 2112 CCATGAGTGTGACAAATGCTGCTGTGTGTGTGTG- AATGTGAGAAAAACACAG 2170

Db 721 CCATGAGTGTGAC- ACATGCTATGGGTCTGNAATGTGTGAAATGTGAGAGACACAG 779

QY 2171 CCTCTCTTTCAGAGAGAGGGGCTGAGGTGCCAGCTGTGTCTGGTTA 2221

Db 780 CCTCTCTTTCAGAGAGAGGGGCTGAGGGGCCACTGTGTCCCGGGTAA 830

RESULT 11  
 BE396312  
 LOCUS 601289231F1 NIH\_MGC\_8 Homo sapiens cDNA clone IMAGE:3619783 5',  
 DEFINITION mRNA sequence.  
 BE396312  
 BE396312.1 GI:9341677  
 EST.  
 Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 839)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: LLC290 row: n column: 08  
 High quality sequence stop: 745.  
 Location/Qualifiers  
 1..839  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3619783"  
 /tissue\_type="Burkitt lymphoma"  
 /lab\_host="DH10B (phage-resistant)"

## FEATURES

source

/clone lib="NIH MGC 8"  
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

Query Match 32.1%; Score 747; DB 10; Length 839;  
 Best Local Similarity 95.4%; Pred. No. 4.6e-158;  
 Matches 802; Conservative 0; Mismatches 35; Indels 4; Gaps 3;

QY 652 TGTGGCTTCGTGGTCATCATGTTTCCCTGCTGGCTGGCTGCTTATCTTCTGAA 711

Db 1 TGTGGCTTCGTGGTCATCATGTTTCCCTGCTGGCTGGCTGCTTATCTTCTGAA 60

QY 712 CTGCACCTCAACTGGCCCATCGAAGCCTTTCCTGCCCTGAGGAAGTCAATTACACGAA 771

Db 61 CTGCACCTCAACTGGCCCATCGAAGCCTTTCCTGCCCTGAGGAAGTCAATTACACGAA 120

QY 772 GAAGATCAAGCTGAGTGGGCTGGCCCTGGACCAAGGTGACAGTGAAGTCTTCTACAC 831

Db 121 GAAGATCAAGCTGAGTGGGCTGGCCCTGGACCAAGGTGACAGTGAAGTCTTCTACAC 180

QY 832 CCATGTGACCAACATGGGCCAGAGGCTCAGCCAGAGGCCCCCGCTGGAGAGCGTTTC 891

Db 181 CCATGTGACCAACATGGGCCAGAGGCTCAGCCAGAGGCCCCCGCTGGAGAGCGTTTC 240

QY 892 GGATGCTTTCATGTCAACCCAGGATGTTTCGGGGCACCCTCAGAAAAACCTTCTGAGAGGTC 951

Db 241 GGATGCTTTCATGTCAACCCAGGATGTTTCGGGGCACCCTCAGAAAAACCTTCTGAGAGGTC 300

QY 952 TGTCCCTTTACGCAAGAGCCTCTGTCCCCCACTTCTGTGGAGCCTCTCACCANTGG 1011

Db 301 TGTCCCTTTACGCAAGAGCCTCTGTCCCCCACTTCTGTGGAGCCTCTCACCANTGG 360

QY 1012 CATGACCCAGCTCGGATCATCTTCTACATGCTGCTGTGAACAGATGCTGGAGTACCT 1071

Db 361 CATGACCCAGCTCGGATCATCTTCTACATGCTGCTGTGAACAGATGCTGGAGTACCT 420

QY 1072 TGTGACTGTGGCCAGGAGCATGAGACAAATGAACAGCAACAAAAGGTGGCAGACAGT 1131

Db 421 TGTGACTGTGGCCAGGAGCATGAGACAAATGAACAGCAACAAAAGGTGGCAGACAGT 480

QY 1132 TGGGTTTACTCTCTCGCTTTCGGGGCCATGAGCTGTTGTGCTTCTCACTGCCCTT 1191

Db 481 TGGGTTTACTCTCTCGCTTTCGGGGCCATGAGCTGTTGTGCTTCTCACTGCCCTT 540

QY 1192 CATTGGCTTACATCATGGACTGGCGATCAAGGACTGCGTGGAGCCCCCAACTCAGGGCAC 1251

Db 541 CATTGGCTTACATCATGGACTGGCGATCAAGGACTGCGTGGAGCCCCCAACTCAGGGCAC 600

QY 1252 TGTCTCGGAGATGCCAGGACGGGTTGCTTACCAAAATCCATCAGACACACGCTACTGCAA 1311

Db 601 TGTCTCGGAGATGCCAGGACGGGTTGCTTACCAAAATCCATCAGACACACGCTACTGCAA 660

QY 1312 GATCAAAAGCTCACAATGCGCTTACCTTCAACCTGACCAACCTGCTGCTTGTGGG 1371

Db 661 GATCAAAAGCTCAACCAATGCTCAGTGGCTTCAACCTGACCAACCTGCTGCTTGTGGG 720

QY 1372 TTTTGGCATCA-CTGTCTCATCAACACTTACACTCCAGTTTGTGACCTTTGCTTCC 1430

Db 721 GTTGGCTTCACTCTCTCAACAACTTACACTCCAGTTTGTGACCTTTGCTTCC- 779

QY 1431 ACACCATTTGTCAGGTTTCTTCCACTCAGCTGTGGAGTCTCTATGTGAGTGTTC 1490

Db 780 -CCCCCTTGTTCGAGGTTTCTTCC-CTCGGCTGTGGAGTCTCTATGTGAGTGTTC 837

QY 1491 C 1491







Best Local Similarity 95.8%; Pred. No. 7.9e-155;		Matches 818; Conservative 0; Mismatches 27; Indels 9; Gaps 6;	
Qy	864	AGAGGCCCCCAGCTGGAGGAGG--TTCCGATGCTTCAT--GTACCCAGGATGTT	919
Db	854	AAAGGCCCCCAGCTGGAGGAGGTTTCGAATGCTTCATTTGCCCCCAAGGATGTT	795
Qy	920	CGGGGACCTTCAGAAACCTTCTCGAGAGTCTGTCCCTTA--CGCAAGAGCCTCTGCT	977
Db	794	CGGGGCTTCAGAAACCTTCTTGAGAGTCTGTCCCTTAACGCAAAAGGCTTCTGTT	735
Qy	978	CCCCCACTTCTCTGCGG--AGCTCTCTCACC--ATGGGATGACCCAGCTCGGATCATCTT	1035
Db	734	CCCCCACTTCTCTGCGGAGCTCTCACCATGGGATGACCCAGCTCGGATCATCTT	675
Qy	1036	CTACATGGCTGCTGCAACAGATGCTGGAGTACCTTGTGACTGGTGGCCAGGAGCATCA	1095
Db	674	CTACATGGCTGCTGCAACAGATGCTGGAGTACCTTGTGACTGGTGGCCAGGAGCATCA	615
Qy	1096	GACAAATGAACAGCAACAAAGGTGCGAGAGCAGTGTGGTTCCTCTCTCTCTCTCGG	1155
Db	614	GACAAATGAACAGCAACAAAGGTGCGAGAGCAGTGTGGTTCCTCTCTCTCTCTCGG	555
Qy	1156	GGCCATGCGAGCTTGTGCTCTCTACCTGCCCCCTCATTTGGCTACATCATGAGCTGGCG	1215
Db	554	GGCCATGCGAGCTTGTGCTCTCTACCTGCCCCCTCATTTGGCTACATCATGAGCTGGCG	495
Qy	1216	GATCAGGAGCTGGTGGAGCGCCCACTCAGGCACTGTCTCGGAGATGCCAGGAGCG	1275
Db	494	GATCAGGAGCTGGTGGAGCGCCCACTCAGGCACTGTCTCGGAGATGCCAGGAGCG	435
Qy	1276	GGTGTGTACCAATCCATCAGACACGCTACTGCAAGATCCAAAGCTCACCATGCCAT	1335
Db	434	GGTGTGTACCAATCCATCAGACACGCTACTGCAAGATCCAAAGCTCACCATGCCAT	375
Qy	1336	CAGTGCCTTACCCCTGACCAACCTGCTGTGTGGTGTGGGATCACTGTCTCATCAA	1395
Db	374	CAGTGCCTTACCCCTGACCAACCTGCTGTGTGGTGTGGGATCACTGTCTCATCAA	315
Qy	1396	CAACTTACACCTCCAGTTGTGACCTTGTCTGACACCATTTTCGAGGTTCTTCCA	1455
Db	314	CAACTTACACCTCCAGTTGTGACCTTGTCTGACACCATTTTCGAGGTTCTTCCA	255
Qy	1456	CTCAGCTGTGGAGTCTCTATGCTGAGTGTTCCTCAACCTTTGGGACGCTGAC	1515
Db	254	CTCAGCTGTGGAGTCTCTATGCTGAGTGTTCCTCAACCTTTGGGACGCTGAC	195
Qy	1516	AGGCTGCGAGTCCCTCATCAGTGTCTGTGTGGCTTGTCTGAGCAGCCACTTTTCATGGC	1575
Db	194	AGGCTGCGAGTCCCTCATCAGTGTCTGTGTGGCTTGTCTGAGCAGCCACTTTTCATGGC	135
Qy	1576	GATGTTGGGAGCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGGCTCTCTGTTATCTC	1635
Db	134	GATGTTGGGAGCCCTGAAAGGAGAGCCCTTCTGGGTGAATCTGGGCTCTCTGTTATCTC	75
Qy	1636	ACTCTCTGGGATTCCTGTGCTTCTTACCTTCTTATACCGTGGCCGCT--CCAGCAGG	1694
Db	74	ACTCTCTGGGATTCCTGTGCTTCTTACCTTCTTATACCGTGGCCGCTTCCAGCAGG	15
Qy	1695	AGTACGCCGCCAAT 1708	
Db	14	AGTACGCCGCCAAT 1	

RESULT 14  
 BG821730  
 LOCUS 60272748F1 NIH\_MGC\_15 Homo sapiens cdna clone IMAGE:4867427 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG821730  
 VERSION BG821730.1 GI:14169317  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

Homo sapiens		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 831)		NIH-MGC http://mgi.nci.nih.gov/.	
AUTHORS		National Institutes of Health, Mammalian Gene Collection (MGC)	
TITLE		Unpublished (1999)	
JOURNAL		Contact: Robert Strausberg, Ph.D.	
COMMENT		Email: cgapbs-r@mail.nih.gov	
		Tissue Procurement: ATCC	
		cDNA Library Preparation: Ling Hong/Rubin Laboratory	
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
		DNA Sequencing by: NIH Intramural Sequencing Center	
		Clone distribution: MGC clone distribution information can be	
		found through the I.M.A.G.E. Consortium/LLNL at:	
		http://image.llnl.gov	
		Plate: LLCMI735 row: o column: 12	
		High quality sequence stop: 790.	
FEATURES		Location/Qualifiers	
source		1..831	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:4867427"	
		/issue_type="adenocarcinoma cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_15"	
		/note="Organ: colon; Vector: pOTB7; Site: 1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"	
ORIGIN		Query Match 31.3%; Score 728.6; DB 12; Length 831;	
		Best Local Similarity 98.2%; Pred. No. 6.7e-154;	
		Matches 779; Conservative 0; Mismatches 9; Indels 5; Gaps 4;	
Qy	1	CCGGGGCTGGAGGGGGCAAGCGGTTCCGAGTGCAAAAGCTGTGTCGCCCGAGCCCTGC	60
Db	44	CCGGGGCTGGAGGGGGCAAGCGGTTCCGAGTGCAAAAGCTGTGTCGCCCGAGCCCTGC	103
Qy	61	GGAGCTCGGGGCGAGCATGGCCCCCAGCTGCAACAGGCGTACCGAGGCGCTGGTGGAT	120
Db	104	GGAGCTCGGGGCGAGCATGGCCCCCAGCTGCAACAGGCGTACCGAGGCGCTGGTGGAT	163
Qy	121	GGCCCTGCAAGCTGTGTGGAGAACTCTTCTCTGCTACTCTCTGGGCTGGGGCTC	180
Db	164	GGCCCTGCAAGCTGTGTGGAGAACTCTTCTCTGCTACTCTCTGGGCTGGGGCTC	223
Qy	181	CCTGTTGATCATTTCTGAAGAACGAGGGCTTTCTATTCCAGCAGTGCCTCAGCTGAGAGCAG	240
Db	224	CCTGTTGATCATTTCTGAAGAACGAGGGCTTTCTATTCCAGCAGTGCCTCAGCTGAGAGCAG	283
Qy	241	GACCAACACCCAGGATGAGCAGGCGAGGTGGCCAGGCTGTGACCAAGGACAGAT	300
Db	284	GACCAACACCCAGGATGAGCAGGCGAGGTGGCCAGGCTGTGACCAAGGACAGAT	343
Qy	301	GCTCAACCTGGGCTTTCACCATTTGTTTCTTCTGCTCAGCGCACCCACCTGCGACTGGG	360
Db	344	GCTCAACCTGGGCTTTCACCATTTGTTTCTTCTGCTCAGCGCACCCCTGCGACTGGG	403
Qy	361	GATCTCTATGAGCCGCTTTGGCCCCCGACCCGCGGCTGGTTGGCAGTGCCTGCTTCAC	420
Db	404	GATCTCTATGAGCCGCTTTGGCCCCCGACCCGCGGCTGGTTGGCAGTGCCTGCTTCAC	463
Qy	421	TGCGTCTGACCCCTCATGGCCCTTGGCCCTCCCGGACGCTGGAGGCTCTGCTCCGTTGAT	480
Db	464	TGCGTCTGACCCCTCATGGCCCTTGGCCCTCCCGGACGCTGGAGGCTCTGCTCCGTTGAT	523

QY 481 ATTCTGGCGCTGCTCCGTAATGGCTTTGGTGGCATCTGCTTAACGTTCACTTCACTCAC 540  
 |||||  
 Db 524 ATTCTGGCGCTGCTCCGTAATGGCTTTGGTGGCATCTGCTTAACGTTCACTTCACTCAC 583  
 |||||  
 QY 541 GCTGCCCAACATGTTTGGGAACCTCGCTGCCATCCAGTTAAATGGCCCTCATGATGGCTCTTA 600  
 |||||  
 Db 584 GCTGCCCAACATGTTTGGGAACCTCGCTGCCATCCAGTTAAATGGCCCTCATGATGGCTCTTA 643  
 |||||  
 QY 601 CGCTCTTCTGTCATATGCTTCCAGGAATCAAGCTGATCTACGATGCGGTGGCCCTT 660  
 |||||  
 Db 644 CGCTCTTCTGTCATATGCTTCCAGGAATCAAGCTGATCTACGATGCGGTGGCCCTT 703  
 |||||  
 QY 661 CGTGGTCATCATGTTCACTGCTGCTGGCTGCCCTTACTTCTGAACTGCACTGCACT 720  
 |||||  
 Db 704 CGTGGTCATCATGTTCACTGCTGCTGGCTGCCCTTACTTCTGAACTGCA-CCT 762  
 |||||  
 QY 721 CAACTGGCCCATCGAAGCCTTTCCTGCCCTGAGGAAGTCAATACAGGAAGATCAA 780  
 |||||  
 Db 763 CAACTGGCCCATCGAAGC--TTTCTGCCCTTAGGAAGTC-ATTACAGAGAGGA-CAA 818  
 |||||  
 QY 781 GCTGAGTGGGCTG 793  
 |||||  
 Db 819 GCTGATGGGTTG 831  
 |||||

## RESULT 15

LOCUS BG823510 726 bp mRNA linear EST 22-MAY-2001  
 DEFINITION 602728885F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:486844 5',  
 mRNA sequence.

ACCESSION BG823510

VERSION BG823510.1 GI:14171097

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 726)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC <http://mgc.nhl.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: NIH Intramural Sequencing Center

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM1738 row: i column: 21

High quality sequence stop: 712.

## FEATURES

source

1..726

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:486844"

/tissue\_type="adenocarcinoma cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 15"

/note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)"

## ORIGIN

Query Match

Best Local Similarity 30.0%; Score 697.8; DB 12; Length 726;

Pred. No. 6e-147;

Matches 708; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
 QY 2 CGGGGCTGGAGGGGGGCAAGCGGTTCCGAGGTTCAAAGCTGTCGCCCGAGCCCTGCG 61  
 |||||  
 Db 2 CGGGGCTGGAGGGGGGCAAGCGGTTCCGAGGTTCAAAGCTGTCGCCCGAGCCCTGCG 61  
 |||||  
 QY 62 GAGCTCGGGGCGACATGCGCCCAACGCTGCAACAGCGCTACCGGAGGCGCTGGTGAATG 121  
 |||||  
 Db 62 GAGCTCGGGGCGACATGCGCCCAACGCTGCAACAGCGCTACCGGAGGCGCTGGTGAATG 121  
 |||||  
 QY 122 GCCTGCAAGGCTGTGTGGAGAACCTCTTCTTCTCTGTCTGTCTGTGGGCTGGGCTCC 181  
 |||||  
 Db 122 GCCTGCAAGGCTGTGTGGAGAACCTCTTCTTCTCTGTCTGTCTGTGGGCTGGGCTCC 181  
 |||||  
 QY 182 CTGTTGATCATTTCTGAAGAACGAGGCTTCTATTCCAGCAGCTGCCAGCTGAGAGCAGC 241  
 |||||  
 Db 182 CTGTTGATCATTTCTGAAGAACGAGGCTTCTATTCCAGCAGCTGCCAGCTGAGAGCAGC 241  
 |||||  
 QY 242 ACCAACACCAACCCAGGATGAGCAGCGAGGTGGCCAGGCTGTGACACGACGAGAGATG 301  
 |||||  
 Db 242 ACCAACACCAACCCAGGATGAGCAGCGAGGTGGCCAGGCTGTGACACGACGAGAGATG 301  
 |||||  
 QY 302 CTCAACCTGGGCTTCAACCATTTGGTTCTTCTGTCTCAGCGCCACACCTGCCACTGGGG 361  
 |||||  
 Db 302 CTCAACCTGGGCTTCAACCATTTGGTTCTTCTGTCTCAGCGCCACACCTGCCACTGGGG 361  
 |||||  
 QY 362 ATCTCTATGACCGCTTTGGCCCCCGAGCCGCTGGGCTGGTGGCAGTGCCTTCACT 421  
 |||||  
 Db 362 ATCTCTATGACCGCTTTGGCCCCCGAGCCGCTGGGCTGGTGGCAGTGCCTTCACT 421  
 |||||  
 QY 422 GCGTCTCGACCCCTCATGGCCCTCGGCTCCCGGACGTTGGAAGCTCTGTCTCCGTTGATA 481  
 |||||  
 Db 422 GCGTCTCGACCCCTCATGGCCCTCGGCTCCCGGACGTTGGAAGCTCTGTCTCCGTTGATA 481  
 |||||  
 QY 482 TTCTGCGCGCTGTCCCTGAATGGCTTTGGTGGCAATCTGCTTAAACGTTCACTCACTACG 541  
 |||||  
 Db 482 TTCTGCGCGCTGTCCCTGAATGGCTTTGGTGGCAATCTGCTTAAACGTTCACTCACTACG 541  
 |||||  
 QY 542 CTGCCCCAACATGTTTGGGAACCTCGCTCCACGTTAATGGCCCTCATGATGGGCTCTTAC 601  
 |||||  
 Db 542 CTGCCCCAACATGTTTGGGAACCTCGCTCCACGTTAATGGCCCTCATGATGGGCTCTTAC 601  
 |||||  
 QY 602 GCCTCTTCTGCAATTAGCTTCCAGGAATCAAGCTGATCTACGATGCGGCTGTGGCCTTC 661  
 |||||  
 Db 602 GCCTCTTCTGCAATTAGCTTCCAGGAATCAAGCTGATCTACGATGCGGCTGTGGCCTTC 661  
 |||||  
 QY 662 GTGTCATCATGTTTCACTGCTGTGGGCTGGCCCTGCTTATCTTTCTGAACTGCACTCCTC 721  
 |||||  
 Db 662 GTGTCATCATGTTTCACTGCTGTGGGCTGGCCCTGCTTATCTTTCTGAACTGCACTCCTC 721  
 |||||  
 QY 722 AACTG 726  
 |||||  
 Db 722 AGATG 726  
 |||||

Search completed: April 10, 2004, 19:48:47

Job time : 7932.59 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using ew model

Run on: April 10, 2004, 14:24:52 ; Search time 8.38845 Seconds  
(without alignments)  
10128.688 Million cell updates/sec

Title: US-09-743-825-10

Perfect score: 20

Sequence: 1 gaccgcagatcttcagc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001as.\*

5: Geneseq2001bs.\*

6: Geneseq2002s.\*

7: Geneseq2003as.\*

8: Geneseq2003bs.\*

9: Geneseq2003cs.\*

10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	3	Aaz50446 Human PB3
2	20	100.0	457	6	Abn93510 Gene #8 u
3	20	100.0	856	4	Aal01556 Human rep
4	20	100.0	856	4	Abi96993 Human tes
5	20	100.0	1737	4	Aal05399 Human rep
6	20	100.0	1737	4	Abi98261 Human tes
7	20	100.0	2093	4	Aah16350 Human CDN
8	20	100.0	2105	7	Acc72138 Human NOV
9	20	100.0	2233	7	Acc20413 DNA encod
10	20	100.0	2284	7	Acc72139 Human NOV
11	20	100.0	2326	3	Aaz50442 Human PB3
12	20	100.0	2326	9	Adb75506 Prostata
13	20	100.0	2373	2	Aav49590 Human liv
14	20	100.0	3160	4	Aal05398 Human rep
15	20	100.0	3160	4	Abi98260 Human tes
16	20	100.0	7709	4	Aal05397 Human tes
17	20	100.0	7709	4	Abi98259 Human tes
18	20	100.0	143899	6	Aal38336 Genomic s
19	15.8	79.0	539	6	Abq98188 Mouse ES
20	15.8	79.0	1232	6	Abk34925 Human CDN
21	15.8	79.0	1656	4	Acc88151 Human CD7
22	15.8	79.0	1665	1	Aan90608 CD7 cDNA
23	15.8	79.0	1665	2	Aaq21168 Human CD7

24	15.8	79.0	1665	2	AAT14708	Aat14708 Human CD7
25	15.8	79.0	1665	2	AAV63446	Aav63446 Human CD7
26	15.8	79.0	1665	2	AAV81204	Aav81204 Human CD7
27	15.8	79.0	1665	3	AAAS0582	Aas0582 Human cel
28	15.8	79.0	1665	4	AAAS03176	Aas03176 Human lym
29	15.8	79.0	3280	6	ABL34271	Abi34271 Human inm
30	15.8	79.0	3280	6	ABL34631	Abi34631 Human met
31	15.4	77.0	349	4	AAK65165	Aak65165 Human inm
32	15.4	77.0	656	8	ADB82432	Adb82432 Human CDN
33	15.4	77.0	658	8	ADB82357	Adb82357 Human CDN
34	15.4	77.0	680	6	ABQ89489	Abq89489 Human pro
35	15.4	77.0	704	8	ADB82873	Adb82873 Human CDN
36	15.4	77.0	737	8	ADB82925	Adb82925 Human CDN
37	15.4	77.0	751	6	ABQ89414	Abq89414 Human pro
38	15.4	77.0	793	6	ABQ89886	Abq89886 Human pro
39	15.4	77.0	797	8	ADB83057	Adb83057 Human CDN
40	15.4	77.0	817	6	ABQ89834	Abq89834 Human pro
41	15.4	77.0	2103	5	AAH67661	Aah67661 C glutam
42	15.4	77.0	2151	7	ACA01132	Aca01132 C. glutam
43	15.4	77.0	349980	5	AAH68532	Aah68532 C glutam
44	15.2	76.0	36	3	AAZ91060	Aaz91060 Primer P5
45	15.2	76.0	36	4	AAD18774	Aad18774 PCR prime

ALIGNMENTS

RESULT 1

AAZ50446

ID AAZ50446 standard; DNA; 20 BP.

AC AAZ50446;

DT 18-MAY-2000 (first entry)

DE Human PB39 specific 3' RACE primer.

PB39; human; prostate cancer; PC; chromosome 11p11.1-11.2; cancer;  
prostate epithelium; splicing mechanism; early diagnosis; progression;  
precancerous cell; metastatic potential; non-neoplastic prostate disease;  
expressed sequence tag; EST; PCR primer; ss.

OS Homo sapiens.

PN WO200005376-A1.

PD 03-FEB-2000.

PF 23-JUL-1999; 99WO-US016831.

XX 24-JUL-1998; 98US-0094137P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Chuquai RF, Cole KA, Liotta LA;

XX WPI; 2000-182700/16.

XX Novel gene which is dysregulated in prostate cancer useful for diagnosing cancer.

XX Claim 5; Page 18; 51pp; English.

XX The present sequence is the human PB39 3' specific RACE primer, from EST  
clone AAR00504. It is used to determine the complete nucleotide sequence  
of PB39 cDNA, isolated from human pancreas cDNA library using RACE. The  
PB39 gene that is dysregulated in prostate cancer has homology to the EST  
AAR00504. PB39 gene is located on chromosome 11p11.1-11.2. Abnormally  
high concentrations of PB39 are found in prostate tissue derived from  
prostate cancer (PC) epithelium. PB39 sequence is useful for detection of  
precancerous or cancer cells in the prostate. PB39 is useful for early  
diagnosis of the progression of prostate cancer, especially in aggressive  
prostate carcinoma. It can also distinguish PC from other non-neoplastic

CC prostate disease. The diagnostic method is selective and specific for  
CC various types of PC and also facilitates identifying prostate cancer of  
CC differing aggressiveness and metastatic potential

XX  
SQ Sequence 20 BP; 6 A; 6 C; 4 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACCGCATAGACTTCTCAGA 20  
|||  
DB 1 GACCGCATAGACTTCTCAGA 20  
|||

RESULT 2  
ID ABN93510/c  
ID ABN93510 standard; DNA; 457 BP.

XX AC ABN93510;

XX DT 13-AUG-2002 (first entry)

XX DE Gene #8 used to diagnose liver cancer.

XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
KW metastatic liver tumour; cytostatic; expression profile; disease state;  
KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX OS Homo sapiens.

XX PN WO200229103-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US030589.

XX PR 02-OCT-2000; 2000US-0237054P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX DR WPI; 2002-426119/45.

XX PT Diagnosing and detecting the progression of liver cancer, hepatocellular  
PT carcinoma or metastatic liver tumor in a patient, involves detecting the  
PT level of expression of two or more genes in a liver tissue sample.

XX PS Claim 1; SEQ ID NO 8; 298pp; English.

XX CC The invention relates to a novel method for diagnosing and detecting the  
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
CC tumour in a patient, and differentiating metastatic liver cancer from  
CC hepatocellular carcinoma in a patient, involving detecting the level of  
CC expression of two or more genes represented in ABN935103-ABN97455 in a  
CC tissue sample. The method of the invention has hepatotropic, and  
CC cytostatic activity. The method is useful for diagnosing and detecting  
CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 457 BP; 117 A; 108 C; 119 G; 113 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 457;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACCGCATAGACTTCTCAGA 20  
|||  
DB 144 GACCGCATAGACTTCTCAGA 125  
|||

RESULT 3

AAAL01556/c

ID AAL01556 standard; cDNA; 856 BP.

XX AC AAL01556;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen cDNA SEQ ID NO: 1557.

XX KW Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ss.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001339.

XX PR 31-JAN-2000; 2000US-0179085P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 30-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 18-AUG-2000; 2000US-0226279P.

XX PR 22-AUG-2000; 2000US-0226681P.

XX PR 22-AUG-2000; 2000US-0227182P.

XX PR 23-AUG-2000; 2000US-0227009P.

XX PR 30-AUG-2000; 2000US-0228924P.

XX PR 01-SEP-2000; 2000US-0229287P.

XX PR 01-SEP-2000; 2000US-0229343P.

XX PR 01-SEP-2000; 2000US-0229344P.

XX PR 05-SEP-2000; 2000US-0229345P.

XX PR 05-SEP-2000; 2000US-0229509P.

XX PR 06-SEP-2000; 2000US-0229513P.

XX PR 06-SEP-2000; 2000US-0230437P.

XX PR 08-SEP-2000; 2000US-0230438P.

XX PR 08-SEP-2000; 2000US-0231242P.

XX PR 08-SEP-2000; 2000US-0231243P.

XX PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232297P.  
 PR 14-SEP-2000; 2000US-0232298P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0232402P.  
 PR 14-SEP-2000; 2000US-0232403P.  
 PR 14-SEP-2000; 2000US-0232404P.  
 PR 14-SEP-2000; 2000US-0232405P.  
 PR 14-SEP-2000; 2000US-0232406P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234224P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 25-SEP-2000; 2000US-0234999P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235835P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236328P.  
 PR 29-SEP-2000; 2000US-0236329P.  
 PR 29-SEP-2000; 2000US-0236330P.  
 PR 29-SEP-2000; 2000US-0236331P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239936P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251899P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI: 2001-465570/50.  
 DR P-PSDB; AM95586.  
 DR  
 XX  
 PT Isolated nucleic acid molecule encoding a reproductive system antigen is  
 PT used in preventing, treating or ameliorating a medical condition.  
 XX  
 PS Claim 1; SEQ ID NO 1557; 1297pp + Sequence Listing; English.  
 XX  
 CC The present invention provides the protein and coding sequences of a  
 CC number of human reproductive system related antigens. These can be used  
 CC in the prevention and treatment of reproductive system disorders,  
 CC including cancer. The present sequence is a coding sequence of the  
 CC invention  
 XX  
 SQ Sequence 856 BP; 209 A; 208 C; 234 G; 203 T; 0 U; 2 Other;  
 Query Match 100.0%; Score 20; DB 4; Length 856;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACCGCATAGACTTCTCAGA 20  
 Db 163 GACCGCATAGACTTCTCAGA 144  
 RESULT 4  
 ABL96993/c  
 ID ABL96993 standard; cDNA; 856 BP.  
 XX  
 AC ABL96993;  
 XX  
 DT 21-JUN-2002 (first entry)  
 XX  
 DE Human testicular antigen encoding cDNA SEQ ID NO: 661.  
 XX  
 KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
 KW reproductive system disorder; urinary system disorder; gene therapy;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disease; infection; cytostatic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155317-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001329.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0218680P.  
PR 11-JUL-2000; 2000US-02117487P.  
PR 11-JUL-2000; 2000US-02117496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225211P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-022579P.  
PR 22-AUG-2000; 2000US-0228273P.  
PR 22-AUG-2000; 2000US-0228681P.  
PR 22-AUG-2000; 2000US-0228686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229511P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231411P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483232/52.  
DR  
XX  
XX  
PT Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
for preventing, diagnosing and/or treating testicular cancer.  
XX  
PS Claim 1; SEQ ID NO 661; 766pp; English.  
XX  
CC The present invention provides the protein and coding sequences of 973  
human testicular antigens, and fragments of their genomic sequences. The

CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a cDNA of the  
CC invention

XX  
SQ Sequence 856 BP; 209 A; 208 C; 234 G; 203 T; 0 U; 2 Other;  
Query Match 100.0%; Score 20; DB 4; Length 856;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGCTTCAG 20  
|||||  
DB 163 GACCGCATAGCTTCAG 144

## RESULT 5

AAL05399/c

ID AAL05399 standard; DNA; 1737 BP.

XX AC AAL05399;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen DNA SEQ ID NO: 8087.

XX KW Human; reproductive system related antigen; reproductive system disorder;  
XX KW cancer; gene therapy; ds.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001339.

XX PR 11-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189074P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220953P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.

PR 23-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0230439P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249298P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251989P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-465570/50.  
 DR  
 XX  
 PT Isolated nucleic acid molecule encoding a reproductive system antigen is  
 PR used in preventing, treating or ameliorating a medical condition.  
 XX  
 PS Disclosure; SEQ ID NO 8087; 1297pp + Sequence Listing; English.  
 XX  
 XX The present invention provides the protein and coding sequences of a  
 CC number of human reproductive system related antigens. These can be used  
 CC in the prevention and treatment of reproductive system disorders,  
 CC including cancer. The present sequence is a genomic sequence encoding a  
 CC protein of the invention  
 XX  
 SQ Sequence 1737 BP; 443 A; 431 C; 458 G; 405 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 4; Length 1737;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACCGCATAGACTTCTCAGA 20  
 Db 1044 GACCGCATAGACTTCTCAGA 1025  
 RESULT 6  
 ABL98261/c  
 ID ABL98261 standard; DNA; 1737 BP.  
 XX  
 AC ABL98261;  
 XX  
 DT 21-JUN-2002 (first entry)  
 XX  
 DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2913.  
 XX  
 KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
 KW reproductive system disorder; urinary system disorder; gene therapy;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disease; infection; cytostatic; gene; ds.  
 XX  
 OS Homo sapiens.

XX WO200155317-A2.  
 PN  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001329.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 22-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225577P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226686P.  
 PR 23-AUG-2000; 2000US-0227182P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.



PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239393P.  
 PR 13-OCT-2000; 2000US-0239393P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241321P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246509P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-483232/52.  
 XX  
 PT Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
 for preventing, diagnosing and/or treating testicular cancer.  
 XX  
 PS Disclosure; SEQ ID NO 2913; 766pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of 973  
 human testicular antigens, and fragments of their genomic sequences. The  
 sequences can be used in the treatment of cardiovascular, urinary system,  
 CC reproductive system, immune, respiratory, neurological and  
 CC gastrointestinal disorders, infections, and particularly cancer,  
 CC especially testicular cancers. The present sequence is a DNA encoding a  
 CC protein fragment of the invention  
 XX  
 SQ Sequence 1737 BP; 443 A; 431 C; 458 G; 405 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 4; Length 1737;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACCGCATAGACTTCTCAGA 20  
 |||||  
 DB 1044 GACCGCATAGACTTCTCAGA 1025  
 RESULT 7  
 AAH16350  
 ID AAH16350 standard; cDNA; 2093 BP.  
 XX  
 AC AAH16350;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:15276.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-00116126.  
 XX  
 PR 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX  
 (HELI-) HELIX RES INST.  
 XX  
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.  
 XX  
 PS Claim 8; SEQ ID NO 15276; 2537pp + Sequence Listing; English.  
 XX  
 CC The present invention describes primer sets for synthesising 5602 full-  
 length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
 CC complementary strand of a polynucleotide which comprises one of the 5602  
 CC nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention

XX Sequence 2093 BP; 429 A; 623 C; 548 G; 493 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 2093;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
 |||||  
 Db 1595 GACCGCATAGACTTCTCAGA 1544

RESULT 8

ACD20413

ID ACC72138 standard; DNA; 2105 BP.

XX ACC72138;

XX 07-JUL-2003 (first entry)

XX Human NOV33a coding sequence.

XX Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; neurotropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; metabolic disorder;  
 KW diabetes; obesity; infection; cachexia; cancer; dyslipidaemia;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; gene; ds.

XX Homo sapiens.

OS WO2003029423-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031358.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327342P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 12-OCT-2001; 2001US-0328056P.

XX 15-OCT-2001; 2001US-0329414P.

XX 17-OCT-2001; 2001US-0330142P.

XX 22-OCT-2001; 2001US-0341058P.

XX 24-OCT-2001; 2001US-0332666P.

XX 24-OCT-2001; 2001US-0343629P.

XX 29-OCT-2001; 2001US-0349575P.

XX 01-NOV-2001; 2001US-0346357P.

XX 12-APR-2002; 2002US-0371972P.

XX 12-APR-2002; 2002US-0371980P.

XX 17-APR-2002; 2002US-0373261P.

XX 19-APR-2002; 2002US-0373805P.

XX 23-APR-2002; 2002US-0374738P.

PR 16-MAY-2002; 2002US-0381101P.  
 PR 17-MAY-2002; 2002US-0381635P.  
 PR 29-MAY-2002; 2002US-0383830P.  
 PR 01-OCT-2002; 2002US-00262839.  
 XX  
 PA (CURA-) CURAGEN CORP.

XX Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;  
 PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
 PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;  
 PI Rothenberg MB, Shinkets RA, Smithson G, Spytek KA, Taupier RJ;  
 PI Vernet CAM, Voss EZ, Zethusen BD, Zhong M;

XX WPI; 2003-381625/36.

DR P-PSDB; ABR58426.

XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
 PT dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.

XX Claim 20; Page 209-210; 487pp; English.

XX The present invention relates to novel human NOV proteins and their  
 CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
 CC proteins are useful in manufacturing a medicament for treating a syndrome  
 CC associated with a human disease. The NOV proteins and coding sequences  
 CC may be used to diagnose, treat or prevent metabolic disorders such as  
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
 CC disorders, haematopoietic disorders and various dyslipidaemias

XX Sequence 2105 BP; 425 A; 629 C; 556 G; 495 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 7; Length 2105;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20

|||||

Db 1582 GACCGCATAGACTTCTCAGA 1601

RESULT 9

ACD20413

ID ACD20413 standard; DNA; 2233 BP.

XX AC

XX ACD20413;

XX 26-AUG-2003 (first entry)

XX DNA encoding human NOV28 protein.

XX Human; NOVX; inflammatory disorder; demyelination disease; stroke;

KW renal disorder; infection; cardiomyopathy; atherosclerosis; acne;

KW hypertension; pancreatitis; Von Hippel-Lindau; endometriosis; fertility;

KW scleroderma; cirrhosis; inflammatory bowel disease; Crohn's disease;

KW haemophilia; autoimmune disease; allergy; AIDS;

KW graft versus host disease; Alzheimer's disease; arthritis; pain;

KW Parkinson's disease; Huntington's disease; obesity; diabetes;

KW hair growth; hair loss; asthma; schizophrenia; glomerulonephritis;

KW lupus erythematosus; psoriasis; antidiabetic; anorectic; metabolic;

KW neurotropic; neuroprotective; cytostatic; antibacterial; virucide;

KW protooncogene; antiarteriosclerotic; hypotensive; cerebroprotective;

KW antiinflammatory; gynaecological; antiinfectivity; dermatological;

KW hepatotropic; haemostatic; immunosuppressive; antiallergic;

KW antiarthritic; anticonvulsant; antiseborrhoeic; antiasthmatic;

KW neuroleptic; anti-HIV; analgesic; nephrotropic; antipsoriatic; gene; ds.

XX Homo sapiens.

OS WO200298917-A2.

XX

XX

XX

XX

PD 12-DEC-2002.  
 XX 12-FEB-2002; 2002WO-US022049.  
 PF 12-FEB-2001; 2001US-0268221P.  
 XX 13-FEB-2001; 2001US-0268496P.  
 PR 14-FEB-2001; 2001US-0268646P.  
 PR 15-FEB-2001; 2001US-0268665P.  
 PR 16-FEB-2001; 2001US-0269136P.  
 PR 16-FEB-2001; 2001US-0269310P.  
 PR 16-FEB-2001; 2001US-0269530P.  
 PR 16-MAR-2001; 2001US-0276405P.  
 PR 16-MAR-2001; 2001US-0276399P.  
 PR 16-MAR-2001; 2001US-0276703P.  
 PR 23-MAR-2001; 2001US-0278199P.  
 PR 28-MAR-2001; 2001US-0279274P.  
 PR 30-MAR-2001; 2001US-0280238P.  
 PR 02-APR-2001; 2001US-0280899P.  
 PR 08-AUG-2001; 2001US-0310797P.  
 PR 14-AUG-2001; 2001US-0312284P.  
 PR 14-SEP-2001; 2001US-0322294P.  
 PR 14-SEP-2001; 2001US-0322295P.  
 PR 18-OCT-2001; 2001US-0330293P.  
 PR 31-OCT-2001; 2001US-0335104P.  
 PR 31-OCT-2001; 2001US-0335109P.  
 PR 21-NOV-2001; 2001US-0332127P.  
 PR 28-NOV-2001; 2001US-0331772P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 XX Guo X, Fernandes E, Li L, Kekuda R, Liu Y, Leite M, Spytek KA;  
 PI Ji W, Casman SJ, Boldog FL, Patturajan M, Vernet CAM, Ballinger RA;  
 PI Malvankar UM, Tchernev VT, Blalock AD, Gusev VY, Rastelli L;  
 PI Mezes PD, Ellerman K, Heyes M, Herrmann JL, Shinkets RA, Ioime N;  
 PI Pena CEA, Shenoy SG, Taupier RJ, Gerlach V, Gorman L;  
 XX WPI; 2003-148650/14.  
 DR P-PSDB; ABO15025.  
 DR  
 XX Novel NOVX polypeptide useful for identifying an agent that binds to the  
 PT polypeptide, and for treating cardiomyopathy, atherosclerosis,  
 PT hypertension, infertility, scleroderma, cirrhosis, and inflammatory bowel  
 PT disease.  
 XX  
 PS Claim 9; Page 217-218; 566pp; English.  
 XX  
 CC The present invention relates to the isolation of novel human  
 CC polypeptides referred to as NOVX (NOVI-NOV37), variants of these  
 CC proteins, and the polynucleotide sequences encoding them. The NOVX  
 CC proteins of the invention share homology to various types of protein  
 CC families such as zinc finger-like proteins, enzymes, receptors, and  
 CC lipoproteins. The sequences of the invention may be useful in the  
 CC manufacture of a medicament for treating a syndrome associated with a  
 CC human disease. For example they can be used to treat inflammatory  
 CC disorders, demyelination disease, renal disorders, infections,  
 CC cardiomyopathy, atherosclerosis, hypertension, stroke, pancreatitis, Von  
 CC Hippel-Lindau, endometriosis, fertility, scleroderma, cirrhosis,  
 CC inflammatory bowel disease, Crohn's disease, haemophilia, autoimmune  
 CC diseases, allergies, graft versus host disease, Alzheimer's disease,  
 CC arthritis, Parkinson's disease, Huntington's disease, AIDS, pain,  
 CC acne, hair growth/loss, asthma, schizophrenia, diabetes, diabetes,  
 CC glomerulonephritis, lupus erythematosus, and psoriasis. ACD20372-ACD20427  
 CC represent DNA sequences encoding the NOVX polypeptides of the invention.  
 CC Note: SEQ ID Nos 113-460 are known sequences used for homology purposes  
 XX  
 SQ Sequence 2233 BP; 458 A; 659 C; 605 G; 511 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 7; Length 2233;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 GACCGCATAGACTTCTCAGA 20  
 |||||||||||||||||||

Db 1654 GACCGCATAGACTTCTCAGA 1673  
 RESULT 10  
 ACC72139  
 ID ACC72139 standard; DNA; 2284 BP.  
 XX  
 AC ACC72139;  
 XX 07-JUL-2003 (first entry)  
 XX  
 DE Human NOV33b coding sequence.  
 XX  
 KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; metabolic disorder;  
 KW diabetes; obesity; infection; cachexia; cancer; dyslipidaemia;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003029423-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 XX 02-OCT-2002; 2002WO-US031358.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327342P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 29-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-034575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 12-APR-2002; 2002US-0371972P.  
 PR 12-APR-2002; 2002US-0371980P.  
 PR 17-APR-2002; 2002US-0373261P.  
 PR 19-APR-2002; 2002US-0373805P.  
 PR 23-APR-2002; 2002US-0374738P.  
 PR 16-MAY-2002; 2002US-0381101P.  
 PR 17-MAY-2002; 2002US-0381635P.  
 PR 29-MAY-2002; 2002US-0383830P.  
 PR 01-OCT-2002; 2002US-00262839.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 PI Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;  
 PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
 PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;  
 PI Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ;  
 PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;  
 XX WPI; 2003-381625/36.  
 DR P-PSDB; ABR58427.  
 DR  
 XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
 PT dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 20; Page 210-211; 487pp; English.  
 XX  
 CC The present invention relates to novel human NOV proteins and their  
 CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
 CC proteins are useful in manufacturing a medicament for treating a syndrome

CC associated with a human disease. The NOV proteins and coding sequences  
 CC may be used to diagnose, treat or prevent metabolic disorders such as  
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
 CC disorders, haematopoietic disorders and various dyslipidaemias  
 XX  
 SQ Sequence 2284 BP; 446 A; 697 C; 605 G; 536 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 7; Length 2284;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACCGCATAGACTTCTCAGA 20  
 |||||  
 Db 1762 GACCGCATAGACTTCTCAGA 1781  
 |||||  
 RESULT 11  
 AA250442  
 ID AA250442 standard; cDNA; 2326 BP.  
 XX  
 AC AA250442;  
 XX  
 DT 18-MAY-2000 (first entry)  
 XX  
 DE Human PB39 cDNA, a gene dysregulated in prostate cancer.  
 XX  
 KW PB39; human; prostate cancer; PC; chromosome 11p11.1-11.2; cancer;  
 KW prostate epithelium; splicing mechanism; early diagnosis; progression;  
 KW precancerous cell; aggressive prostate carcinoma; metastatic potential;  
 KW non-neoplastic prostate disease; expressed sequence tag; EST; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT 5'UTR  
 FT 1. .76  
 FT /tag= a  
 FT misc\_signal  
 FT 71. .80  
 FT /tag= b  
 FT /label= Kozak\_consensus\_sequence  
 FT CDS  
 FT 77. .1756  
 FT /tag= c  
 FT /product= "Human PB39 protein"  
 FT /note= "Over-expressed in prostate cancer"  
 FT 3'UTR  
 FT 1757. .2326  
 FT /tag= d  
 FT  
 FN WO200005376-A1.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 23-JUL-1999; 99WO-US016831.  
 XX  
 PR 24-JUL-1998; 98US-0094137P.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Chuquai RP, Cole KA, Liotta LA;  
 XX  
 DR WPI; 2000-182700/16.  
 DR P-PSDB; AAY44897.  
 XX  
 PT Novel gene which is dysregulated in prostate cancer useful for diagnosing  
 PT cancer.  
 XX  
 PS Claim 1; Page 37-40; 51pp; English.  
 XX  
 CC The present sequence is the human PB39 cDNA, a gene that is dysregulated  
 CC in prostate cancer and is isolated from human pancreas cDNA library. It  
 CC has homology to an expressed sequence tag (EST) AA200504. It is mapped to  
 CC chromosome 11p11.1-11.2. A variant of PB39 results from an alternative  
 CC RNA splicing mechanism, yielding a larger transcript (5kb). Abnormally  
 CC high concentrations of PB39 are found in prostate tissue derived from

CC prostate cancer (PC) epithelium. PB39 sequence is useful for detection of  
 CC precancerous or cancer cells in the prostate. PB39 is useful for early  
 CC diagnosis of the progression of prostate cancer, especially in aggressive  
 CC prostate carcinoma. It can also distinguish PC from other non-neoplastic  
 CC prostate disease. The diagnostic method is selective and specific for  
 CC various types of PC and also facilitates identifying prostate cancer of  
 CC differing aggressiveness and metastatic potential  
 XX  
 SQ Sequence 2326 BP; 475 A; 690 C; 625 G; 536 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 3; Length 2326;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACCGCATAGACTTCTCAGA 20  
 |||||  
 Db 1747 GACCGCATAGACTTCTCAGA 1766  
 |||||  
 RESULT 12  
 ADB75506  
 ID ADB75506 standard; cDNA; 2326 BP.  
 XX  
 AC ADB75506;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Prostate cancer marker cDNA.  
 XX  
 KW Prostate; cancer; cytostatic; gene therapy; marker; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003009814-A2.  
 XX  
 PD 06-FEB-2003.  
 XX  
 PF 25-JUL-2002; 2002WO-US023913.  
 XX  
 PR 25-JUL-2001; 2001US-0307982P.  
 PR 22-AUG-2001; 2001US-0314356P.  
 PR 25-SEP-2001; 2001US-0325020P.  
 PR 12-DEC-2001; 2001US-0341746P.  
 PR 05-MAR-2002; 2002US-0362158P.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbatcheva B;  
 PI Hoersh S, Kamatkar S, Wonsay AM, Glatt K, Zhao X, Anderson D;  
 XX  
 DR WPI; 2003-248033/24.  
 XX  
 FT New nucleic acid molecule, useful for diagnosing or treating prostate  
 FT cancer.  
 XX  
 PS Disclosure; SEQ ID NO 330; 99pp; English.  
 XX  
 CC The invention relates to newly discovered cancer markers associated with  
 CC the cancerous state of prostate cells. Also disclosed is a method of  
 CC assessing whether a patient is afflicted with prostate cancer. The method  
 CC of the invention involves assessing whether a patient is afflicted with  
 CC prostate cancer by comparing the level of expression of a marker in a  
 CC patient sample and the normal level of expression of the marker in a  
 CC control non-prostate cancer sample, where a significant increase in the  
 CC level of expression of the marker in the patient sample and the normal  
 CC level indicates that the patient is afflicted with prostate cancer.  
 CC Nucleic acids of the invention are useful for diagnosing or treating  
 CC prostate cancer, and may be useful in gene therapy. Sequences given in  
 CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

SQ Sequence 2326 BP; 475 A; 690 C; 625 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 9; Length 2326;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1747 GACCGCATAGACTTCTCAGA 1766

## RESULT 13

AAV49590  
ID AAV49590 standard; cDNA to mRNA; 2373 BP.

XX AAV49590;

DT 21-OCT-1998 (first entry)

XX Human liver cell clone HF10302 cDNA #1.

XX Transmembrane domain; human; nutrition; cytokine; cell proliferation;  
KW differentiation; immune system; stimulator; suppressor; regulator;  
KW haematopoiesis; activin; inhibitor; chemotactic; chemokinetic; receptor;  
KW haemostatic; thrombolytic; ligand; anti-inflammatory; tumour; ds.

XX Homo sapiens.

PH Key Location/Qualifiers  
FT CDS 134..1813  
/\*tag= a

FT /product= "transmembrane domain containing protein"

XX WO9821328-A2.

PD 22-MAY-1998.

XX 07-NOV-1997; 97WO-JP004056.

XX 13-NOV-1996; 96JP-00301429.

XX (SAGA ) SAGAMI CHEM RES CENTRE.  
PA (PROT-) PROTEGENE INC.

XX Kato S, Sekine S, Yamaguchi T, Kobayashi M;

XX WPI; 1998-297932/26.  
DR P-PSDB; AAW64554.

XX Human protein having transmembrane domain - useful for, e.g. research and nutrition.

PS Claim 4; Page 166-169; 205pp; English.

XX AAV49550-V49599 are cDNA sequences which encode human proteins containing a transmembrane domain. These proteins can be used for, e.g. research and nutrition, and may have cytokine and cell proliferation/differentiation, immune stimulating/suppressing, haematopoiesis regulating, tissue growth, activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic, CC receptor/ligand, anti-inflammatory or tumour inhibition activity

SQ Sequence 2373 BP; 469 A; 716 C; 645 G; 543 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 2373;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1804 GACCGCATAGACTTCTCAGA 1823

## RESULT 14

AA05398/c  
ID AAL05398 standard; DNA; 3160 BP.

XX AAL05398;

DT 21-NOV-2001 (first entry)

XX Human reproductive system related antigen DNA SEQ ID NO: 8086.

XX Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ds.

XX Homo sapiens.

XX WO200155320-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US001339.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216800P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.



PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 14-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231142P.  
PR 08-SEP-2000; 2000US-0231143P.  
PR 08-SEP-2000; 2000US-0231144P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234397P.  
PR 25-SEP-2000; 2000US-0234398P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-483232/52.  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
PT for preventing, diagnosing and/or treating testicular cancer.  
XX  
XX Disclosure; SEQ ID NO 2912; 766pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention  
XX  
XX Sequence 3160 BP; 861 A; 761 C; 855 G; 683 T; 0 U; 0 Other;  
SQ

Query Match 100.0%; Score 20; DB 4; Length 3160;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
 |||||  
 Db 1045 GACCGCATAGACTTCTCAGA 1026

Search completed: April 10, 2004, 20:05:51  
 Job time : 10.3884 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 14:52:03 ; Search time 1.85015 Seconds  
(without alignments)  
5998.994 Million cell updates/sec

Title: US-09-743-825-10  
Perfect score: 20  
Sequence: 1 gaccgatagacttctcaga 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*
  - 2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*
  - 3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*
  - 4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*
  - 5: /cgn2\_6/ptodata/2/ina/PCRUS COMB.seq.\*
  - 6: /cgn2\_6/ptodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	15.8	79.0	246	4	US-09-489-039A-4912
C 2	15.8	79.0	972	4	US-09-252-991A-15702
C 3	15.8	79.0	975	4	US-09-252-991A-15674
C 4	15.8	79.0	1485	4	US-09-252-991A-15656
C 5	15.8	79.0	1656	4	US-09-023-655-1260
C 6	15.2	76.0	36	3	US-08-213-741-3
C 7	15.2	76.0	36	3	US-08-522-336-3
C 8	15.2	76.0	2343	2	US-09-018-760-3
C 9	15.2	76.0	2868	4	US-08-780-562-8
C 10	15.2	76.0	3097	2	US-08-599-4558-1
C 11	15.2	76.0	3097	3	US-09-069-7818-1
C 12	15.2	76.0	3097	4	US-09-137-132-1
C 13	15.2	76.0	3097	4	US-08-864-564A-1
C 14	15.2	76.0	3097	4	US-09-094-410-1
C 15	15.2	76.0	3097	4	US-08-708-1230-1
C 16	15.2	76.0	3097	4	US-08-583-153A-1
C 17	15.2	76.0	3097	4	US-08-570-1420-1
C 18	15.2	76.0	3097	4	US-08-638-524B-1
C 19	15.2	76.0	3854	2	US-08-599-4558-42
C 20	15.2	76.0	3854	3	US-09-069-7818-42
C 21	15.2	76.0	3854	4	US-09-137-132-42
C 22	15.2	76.0	3854	4	US-08-864-564A-42
C 23	15.2	76.0	3854	4	US-09-094-410-42
C 24	15.2	76.0	3854	4	US-08-708-1230-42
C 25	15.2	76.0	3854	4	US-08-638-524B-42
C 26	14.8	74.0	792	1	US-08-565-386-5
C 27	14.8	74.0	1437	4	US-09-252-991A-3858

C 28	14.8	74.0	1545	4	US-09-252-991A-3767	Sequence 3767, Ap
C 29	14.8	74.0	2499	4	US-09-252-991A-2681	Sequence 2681, Ap
C 30	14.8	74.0	2544	4	US-09-252-991A-2775	Sequence 2775, Ap
C 31	14.8	74.0	2891	4	US-09-252-991A-3187	Sequence 3187, Ap
C 32	14.8	74.0	3527	2	US-08-909-965C-7	Sequence 7, Appli
C 33	14.8	74.0	4529	1	US-08-565-386-1	Sequence 1, Appli
C 34	14.8	74.0	519	4	US-09-107-532A-3244	Sequence 3244, Ap
C 35	14.2	71.0	29	2	US-09-014-969-28	Sequence 28, Appl
C 36	14.2	71.0	66	4	US-10-043-511-5	Sequence 5, Appli
C 37	14.2	71.0	453	2	US-08-387-942C-38	Sequence 38, Appl
C 38	14.2	71.0	496	4	US-09-621-976-2087	Sequence 2087, Ap
C 39	14.2	71.0	596	3	US-09-328-111-576	Sequence 576, App
C 40	14.2	71.0	816	4	US-09-543-681A-2321	Sequence 2321, App
C 41	14.2	71.0	1630	4	US-09-023-655-937	Sequence 937, App
C 42	14.2	71.0	2067	4	US-09-016-434-1306	Sequence 1306, Ap
C 43	14.2	71.0	2277	4	US-09-495-050A-155	Sequence 155, App
C 44	14.2	71.0	2362	4	US-09-620-312D-273	Sequence 273, App
C 45	14.2	71.0	2447	2	US-09-014-969-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-09-489-039A-4912/c  
; Sequence 4912, Application US/09489039A  
; Patent No. 6610836

GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 4912

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

; US-09-489-039A-4912

Query Match 79.0%; Score 15.8; DB 4; Length 246;

Best Local Similarity 89.5%; Pred. No. 25;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAG 19

Db 49 GACCGCATAGACTTCTCAG 31

RESULT 2

US-09-252-991A-15702/c

; Sequence 15702, Application US/09252991A

; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 15702

; LENGTH: 972

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; FEATURE:

; NAME/KEY: unsure  
; LOCATION: (669)  
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.  
US-09-252-991A-15702

Query Match 79.0%; Score 15.8; DB 4; Length 972;  
Best Local Similarity 89.5%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAG 19  
| | | | | | | | | | | | | | | | | | | | | |  
DB 417 GACCGCATAGACTTCGCTG 399

RESULT 3  
US-09-252-991A-15674/c  
; Sequence 15674, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 15674  
; LENGTH: 975  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (571)  
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.  
US-09-252-991A-15674

Query Match 79.0%; Score 15.8; DB 4; Length 975;  
Best Local Similarity 89.5%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAG 19  
| | | | | | | | | | | | | | | | | | | | | |  
DB 319 GACCGCATAGACTTCGCTG 301

RESULT 4  
US-09-252-991A-15656  
; Sequence 15656, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 15656  
; LENGTH: 1485  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (415)  
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.

US-09-252-991A-15656

Query Match 79.0%; Score 15.8; DB 4; Length 1485;  
Best Local Similarity 89.5%; Pred. No. 35;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAG 19  
| | | | | | | | | | | | | | | | | | | | | |  
DB 667 GACCGCATAGACTTCGCTG 685

RESULT 5  
US-09-023-655-1260  
; Sequence 1260, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1260:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1656 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g29819  
US-09-023-655-1260

Query Match 79.0%; Score 15.8; DB 4; Length 1656;  
Best Local Similarity 89.5%; Pred. No. 36;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAG 19  
| | | | | | | | | | | | | | | | | | | | | |  
DB 743 GCGCGCATCGACTTCTCAG 761

RESULT 6  
US-08-213-741-3  
; Sequence 3, Application US/08213741  
; Patent No. 6027722

GENERAL INFORMATION: Clague P.  
APPLICANT: Hodgson, Clague P.  
TITLE OF INVENTION: Improved vectors for gene therapy  
FILE REFERENCE: 518.008US1  
CURRENT APPLICATION NUMBER: US/08/213,741  
CURRENT FILING DATE: 1994-03-14  
EARLIER APPLICATION NUMBER: US 08/194,208  
EARLIER FILING DATE: 1994-02-07  
EARLIER APPLICATION NUMBER: US 08/130,638  
EARLIER FILING DATE: 1993-10-01  
EARLIER APPLICATION NUMBER: US 08/097,721  
EARLIER FILING DATE: 1993-07-26  
EARLIER APPLICATION NUMBER: US 08/060,568  
EARLIER FILING DATE: 1993-05-12  
EARLIER APPLICATION NUMBER: US 08/030,766  
EARLIER FILING DATE: 1993-03-12  
EARLIER APPLICATION NUMBER: US 07/968,259  
EARLIER FILING DATE: 1992-10-29  
EARLIER APPLICATION NUMBER: US 07/603,635  
EARLIER FILING DATE: 1990-10-25  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 36  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: An oligonucleotide  
US-08-213-741-3

Query Match 76.0%; Score 15.2; DB 3; Length 36;  
Best Local Similarity 85.0%; Pred. No. 38;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
DB 6 GCGCCGATAGACTTCTGAA 25

RESULT 7  
US-08-522-336-3  
Sequence 3, Application US/08522336A  
Patent No. 6287863  
GENERAL INFORMATION:  
APPLICANT: Hodgson, Clague P.  
TITLE OF INVENTION: Vectors for gene transfer  
FILE REFERENCE: 518.001US1  
CURRENT APPLICATION NUMBER: US/08/522,336A  
CURRENT FILING DATE: 1995-11-09  
EARLIER APPLICATION NUMBER: PCT/US94/02752  
EARLIER FILING DATE: 1994-03-14  
EARLIER APPLICATION NUMBER: US 08/194,208  
EARLIER FILING DATE: 1994-02-07  
EARLIER APPLICATION NUMBER: US 08/130,638  
EARLIER FILING DATE: 1993-10-01  
EARLIER APPLICATION NUMBER: US 08/097,721  
EARLIER FILING DATE: 1993-07-26  
EARLIER APPLICATION NUMBER: US 08/060,568  
EARLIER FILING DATE: 1993-05-12  
EARLIER APPLICATION NUMBER: US 08/030,766  
EARLIER FILING DATE: 1993-03-12  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 36  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: An oligonucleotide  
US-08-522-336-3

Query Match 76.0%; Score 15.2; DB 3; Length 36;  
Best Local Similarity 85.0%; Pred. No. 38;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GACCGCATAGACTTCTCAGA 20  
DB 6 GCGCCGATAGACTTCTGAA 25  
RESULT 8  
US-09-018-760-3  
Sequence 3, Application US/09018760  
Patent No. 5981249  
GENERAL INFORMATION:  
APPLICANT: Shi, Qinwei  
APPLICANT: Tobias, Rowel  
TITLE OF INVENTION: A SINGLE CHAIN CREATINE KINASE MB  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/018,760  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1112-1-050  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2343 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
US-09-018-760-3

Query Match 76.0%; Score 15.2; DB 2; Length 2343;  
Best Local Similarity 85.0%; Pred. No. 79;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
DB 2215 GACCGCTGGGCTTCTCAGA 2234

RESULT 9  
US-08-780-562-8/c  
Sequence 8, Application US/08780562  
Patent No. 6541604  
GENERAL INFORMATION:  
APPLICANT: Matthews, William  
APPLICANT: Bennett, Brian  
TITLE OF INVENTION: WSX RECEPTOR  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California

```

; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/780,562
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585005
; FILING DATE: 01/08/97
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/
; FILING DATE: 01/08/97
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0986R1
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2868 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-780-562-8

```

Query Match 76.0%; Score 15.2; DB 4; Length 2868;  
 Best Local Similarity 85.0%; Pred. No. 82;  
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy 1 GACCCATAGACTTCTCAGA 20
Db 2338 GACCACATAGACTGCACAGA 2319

```

```

RESULT 10
US-08-599-455B-1/c
; Sequence 1, Application US/08599455B
; Patent No. 5972621
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpepper, Janice A.
; TITLE OF INVENTION: METHODS OF IDENTIFYING COMPOUNDS THAT
; MODULATE BODY WEIGHT USING THE OB RECEPTOR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,455B
; FILING DATE: 22-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995

```

```

; APPLICATION NUMBER: 08/569,485
; FILING DATE: 08-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3097 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...2742
; US-08-599-455B-1

Query Match 76.0%; Score 15.2; DB 2; Length 3097;
Best Local Similarity 85.0%; Pred. No. 83;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACCCATAGACTTCTCAGA 20
Db 1881 GACCACATAGACTGCACAGA 1862

RESULT 11
US-09-069-781B-1/c
; Sequence 1, Application US/09069781B
; Patent No. 6287782
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpepper, Janice A.
; APPLICANT: White, David W.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
; THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
; INCLUDING OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,781B
; FILING DATE: 29-APRIL-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/864,564
; FILING DATE: 28-MAY-1997
; APPLICATION NUMBER: US 08/708,123
; FILING DATE: 03-SEP-1996
; APPLICATION NUMBER: US 08/638,524
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: US 08/599,455
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: US 08/583,153

```

;; FILING DATE: 28-DEC-1995  
;; APPLICATION NUMBER: US 08/570,142  
;; FILING DATE: 11-DEC-1995  
;; APPLICATION NUMBER: US 08/569,485  
;; FILING DATE: 08-DEC-1995  
;; APPLICATION NUMBER: US 08/566,622  
;; FILING DATE: 04-DEC-1995  
;; APPLICATION NUMBER: US 08/562,663  
;; FILING DATE: 27-NOV-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Meiklejohn, Ph.D., Anita L.  
;; REGISTRATION NUMBER: 35,283  
;; REFERENCE/DOCKET NUMBER: 07334/082001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 542-5070  
;; TELEFAX: (617) 542-8906  
;; TELEX: 200154  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 3097 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: Coding Sequence  
;; LOCATION: 61...2742  
US-09-069-781B-1

Query Match 76.0%; Score 15.2; DB 3; Length 3097;  
Best Local Similarity 85.0%; Pred. No. 83;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1881 GACCACATAGACTGCACAGA 1862

RESULT 12  
US-09-137-132-1/c  
; Sequence 1, Application US/09137132  
; Patent No. 6380363  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; APPLICANT: Tepper, Robert I.  
; APPLICANT: Culpepper, Janice A.  
; APPLICANT: White, David W.  
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR  
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,  
; TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/137,132  
; FILING DATE: 18-AUG-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/864,564  
; FILING DATE: 28-MAY-1997  
; APPLICATION NUMBER: 08/708,123  
; FILING DATE: 03-SEP-1996  
; APPLICATION NUMBER: 08/638,524  
; FILING DATE: 26-APR-1996

;; APPLICATION NUMBER: 08/599,455  
;; FILING DATE: 22-JAN-1996  
;; APPLICATION NUMBER: 08/583,153  
;; FILING DATE: 28-DEC-1995  
;; APPLICATION NUMBER: 08/570,142  
;; FILING DATE: 11-DEC-1995  
;; APPLICATION NUMBER: 08/569,485  
;; FILING DATE: 08-DEC-1995  
;; APPLICATION NUMBER: 08/566,622  
;; FILING DATE: 04-DEC-1995  
;; APPLICATION NUMBER: 08/562,663  
;; FILING DATE: 27-NOV-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Meiklejohn, Ph.D., Anita L.  
;; REGISTRATION NUMBER: 35,283  
;; REFERENCE/DOCKET NUMBER: 07334/019004  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-542-5070  
;; TELEFAX: 617-542-8906  
;; TELEX: 200154

;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 3097 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: Coding Sequence  
;; LOCATION: 61...2742  
US-09-137-132-1

Query Match 76.0%; Score 15.2; DB 4; Length 3097;  
Best Local Similarity 85.0%; Pred. No. 83;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1881 GACCACATAGACTGCACAGA 1862

RESULT 13  
US-08-864-564A-1/c  
; Sequence 1, Application US/08864564A  
; Patent No. 6395498  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; APPLICANT: Tepper, Robert I.  
; APPLICANT: Culpepper, Janice A.  
; APPLICANT: White, David W.  
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR  
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,  
; TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/864,564A  
; FILING DATE: 28-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/708,123  
; FILING DATE: 03-SEP-1996  
; APPLICATION NUMBER: 08/638,524  
; FILING DATE: 26-APR-1996

;; FILING DATE: 26-APR-1996  
;; APPLICATION NUMBER: 08/599,455  
;; FILING DATE: 22-JAN-1996  
;; APPLICATION NUMBER: 08/583,153  
;; FILING DATE: 28-DEC-1995  
;; APPLICATION NUMBER: 08/570,142  
;; FILING DATE: 11-DEC-1995  
;; APPLICATION NUMBER: 08/569,485  
;; FILING DATE: 08-DEC-1995  
;; APPLICATION NUMBER: 08/566,622  
;; FILING DATE: 04-DEC-1995  
;; APPLICATION NUMBER: 08/562,663  
;; FILING DATE: 27-NOV-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Meiklejohn, Ph.D., Anita L.  
;; REGISTRATION NUMBER: 35,283  
;; REFERENCE/DOCKET NUMBER: 07334/019002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-542-5070  
;; TELEFAX: 617-542-8906  
;; TELEX: 200154  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 3097 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: Coding Sequence  
;; LOCATION: 61...2742  
US-08-864-564A-1

Query Match 76.0%; Score 15.2; DB 4; Length 3097;  
Best Local Similarity 85.0%; Pred. No. 83;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1881 GACCACATAGACTGCACAGA 1862

RESULT 14  
US-09-094-410-1/c  
; Sequence 1, Application US/09094410  
; Patent No. 6403552  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; APPLICANT: Tepper, Robert I.  
; APPLICANT: Culpepper, Janice A.  
; APPLICANT: White, David W.  
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR  
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/094,410  
; FILING DATE: 09-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/864,564  
; FILING DATE: 28-MAY-1997

;; APPLICATION NUMBER: 08/708,123  
;; FILING DATE: 03-SEP-1996  
;; APPLICATION NUMBER: 08/638,524  
;; FILING DATE: 26-APR-1996  
;; APPLICATION NUMBER: 08/599,455  
;; FILING DATE: 22-JAN-1996  
;; APPLICATION NUMBER: 08/583,153  
;; FILING DATE: 28-DEC-1995  
;; APPLICATION NUMBER: 08/570,142  
;; FILING DATE: 11-DEC-1995  
;; APPLICATION NUMBER: 08/569,485  
;; FILING DATE: 08-DEC-1995  
;; APPLICATION NUMBER: 08/566,622  
;; FILING DATE: 04-DEC-1995  
;; APPLICATION NUMBER: 08/562,663  
;; FILING DATE: 27-NOV-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Meiklejohn, Ph.D., Anita L.  
;; REGISTRATION NUMBER: 35,283  
;; REFERENCE/DOCKET NUMBER: 07334/019003  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-542-5070  
;; TELEFAX: 617-542-8906  
;; TELEX: 200154  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 3097 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: Coding Sequence  
;; LOCATION: 61...2742  
US-09-094-410-1

Query Match 76.0%; Score 15.2; DB 4; Length 3097;  
Best Local Similarity 85.0%; Pred. No. 83;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1881 GACCACATAGACTGCACAGA 1862

RESULT 15  
US-08-708-123D-1/c  
; Sequence 1, Application US/08708123D  
; Patent No. 6482927  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; APPLICANT: Tepper, Robert I.  
; APPLICANT: Culpepper, Janice A.  
; APPLICANT: White, David W.  
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR  
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/708,123D  
; FILING DATE: 03-SEP-1996

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/638,524
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: 08/599,455
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: 08/569,485
; FILING DATE: 08-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/019001
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3097 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...2742
; US-08-708-123D-1

```

```

Query Match      76.0%; Score 15.2; DB 4; Length 3097;
Best Local Similarity 85.0%; Pred. No. 83;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GACCGCATAGACTTCTCAGA 20
      |||||
Db      1881 GACCACATAGACTGCACAGA 1862

```

Search completed: April 10, 2004, 20:25:41  
Job time : 2.85015 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: April 10, 2004, 14:29:57 ; Search time 7.87777 Seconds  
(without alignments)  
9524.167 Million cell updates/sec

Title: US-09-743-825-10

Perfect score: 20

Sequence: 1 gaccgcagatgactctcaga 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2475585 seqs, 1875730760 residues

Total number of hits satisfying chosen parameters: 4951170

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	366	14	US-10-066-543-1782
2	20	100.0	457	9	US-09-880-107-8
3	20	100.0	856	10	US-09-764-891-1557
4	20	100.0	1737	10	US-09-764-891-8087
5	20	100.0	2105	12	US-10-262-839-127
6	20	100.0	2233	15	US-10-074-978A-83
7	20	100.0	2284	12	US-10-262-839-129
8	20	100.0	2326	12	US-10-342-887-932
9	20	100.0	2326	14	US-10-205-823-330
10	20	100.0	2373	10	US-09-284-320-71
11	20	100.0	3160	10	US-09-764-891-8086
12	20	100.0	7709	10	US-09-764-891-8085
13	20	100.0	143899	10	US-09-972-546-15
14	16.4	82.0	642	15	US-10-027-632-223788
15	15.8	79.0	539	9	US-09-728-446-1456

16	15.8	79.0	545	15	US-10-027-632-216889	Sequence 216889,
17	15.8	79.0	561	15	US-10-369-493-46556	Sequence 46556, A
18	15.8	79.0	602	15	US-10-027-632-12267	Sequence 12267, A
19	15.8	79.0	1232	9	US-09-822-849A-63	Sequence 63, Appl
20	15.8	79.0	1656	9	US-09-997-165-1	Sequence 1, Appl
21	15.8	79.0	3280	14	US-10-311-455-2244	Sequence 2244, Ap
22	15.8	79.0	3280	14	US-10-240-485-184	Sequence 184, App
23	15.4	77.0	656	15	US-10-012-697-744	Sequence 744, App
24	15.4	77.0	658	15	US-10-012-697-669	Sequence 669, App
25	15.4	77.0	663	15	US-10-027-632-243713	Sequence 243713,
26	15.4	77.0	663	15	US-10-027-632-243714	Sequence 243714,
27	15.4	77.0	704	15	US-10-012-697-1086	Sequence 1086, Ap
28	15.4	77.0	737	15	US-10-012-697-1138	Sequence 1138, Ap
29	15.4	77.0	797	15	US-10-012-697-1270	Sequence 1270, Ap
30	15.4	77.0	2103	9	US-09-738-626-2696	Sequence 2696, Ap
31	15.4	77.0	3309400	9	US-09-738-626-1	Sequence 1, Appli
32	15.2	76.0	238	9	US-09-960-352-8845	Sequence 8845, Ap
33	15.2	76.0	403	9	US-09-925-299-308	Sequence 308, App
34	15.2	76.0	403	10	US-09-925-299-308	Sequence 308, App
35	15.2	76.0	537	14	US-10-029-386-3025	Sequence 3025, Ap
36	15.2	76.0	537	14	US-10-029-386-12779	Sequence 12779, A
37	15.2	76.0	675	15	US-10-369-493-42358	Sequence 42358, A
38	15.2	76.0	776	15	US-10-027-632-28529	Sequence 28529, A
39	15.2	76.0	1143	14	US-10-412-222-4	Sequence 4, Appli
40	15.2	76.0	1143	14	US-10-412-223-4	Sequence 4, Appli
41	15.2	76.0	1182	12	US-10-282-122A-15307	Sequence 15307, A
42	15.2	76.0	1360	9	US-09-880-107-2269	Sequence 2269, Ap
43	15.2	76.0	1360	10	US-09-873-367C-707	Sequence 707, App
44	15.2	76.0	1367	15	US-10-027-632-252470	Sequence 252470,
45	15.2	76.0	1390	15	US-10-264-049-261	Sequence 261, App

ALIGNMENTS

RESULT 1

US-10-066-543-1782  
; Sequence 1782, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jlang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1782  
; LENGTH: 366  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-066-543-1782

Query Match 100.0%; Score 20; DB 14; Length 366;  
Best Local Similarity 100.0%; Pred. No. 0.97;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20

Db 50 GACCGCATAGACTTCTCAGA 69

## RESULT 2

US-09-880-107-8/c  
; Sequence 8, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 457  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA004521  
US-09-880-107-8

Query Match 100.0%; Score 20; DB 9; Length 457;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
Db 144 GACCGCATAGACTTCTCAGA 125

## RESULT 3

US-09-764-891-1557/c  
; Sequence 1557, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1557  
; LENGTH: 856  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (2)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-891-1557

Query Match 100.0%; Score 20; DB 10; Length 856;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
Db 163 GACCGCATAGACTTCTCAGA 144

## RESULT 4

US-09-764-891-8087/c  
; Sequence 8087, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8087  
; LENGTH: 1737  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-891-8087

Query Match 100.0%; Score 20; DB 10; Length 1737;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
Db 1044 GACCGCATAGACTTCTCAGA 1025

## RESULT 5

US-10-262-839-127  
; Sequence 127, Application US/10262839  
; Publication No. US20040038877A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook, John,  
; APPLICANT: Anderson, David W.,  
; APPLICANT: Boldog, Ferenc,  
; APPLICANT: Burgess, Catherine,  
; APPLICANT: Catteston, Elina,  
; APPLICANT: Edinger, Shlomit,  
; APPLICANT: Ellerman, Karen,  
; APPLICANT: Gerlach, Valerie,  
; APPLICANT: Gorman, Linda,  
; APPLICANT: Guo, Xiaojia,  
; APPLICANT: Ji, Weizhen,  
; APPLICANT: Kekuda, Ramesh,  
; APPLICANT: Leach, Martin,  
; APPLICANT: Li, Li,  
; APPLICANT: Miller, Charles,  
; APPLICANT: Patturajan, Meera,  
; APPLICANT: Reiser, Daniel,  
; APPLICANT: Rothenberg, Mark,  
; APPLICANT: Shimkets, Richard,  
; APPLICANT: Smithson, Glennda,  
; APPLICANT: Spytek, Kimberly,  
; APPLICANT: Taupier, Raymond, Jr.,  
; APPLICANT: Vernet, Corine,  
; APPLICANT: Voss, Edward,  
; APPLICANT: Zernusen, Brian,  
; APPLICANT: Zhong, Mei

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-462A  
; CURRENT APPLICATION NUMBER: US/10/262,839  
; CURRENT FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,101  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/371,972  
; PRIOR FILING DATE: 2002-04-12  
; PRIOR APPLICATION NUMBER: 60/327,342  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 60/328,044

```

; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 127
; LENGTH: 2105
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1588)
US-10-262-839-127

Query Match      100.0%; Score 20; DB 12; Length 2105;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACCGCATAGACTTCTCAGA 20
Db      1582 GACCGCATAGACTTCTCAGA 1601

RESULT 6
US-10-074-978A-83
; Sequence 83, Application US/10074978A
; Publication No. US20040010119A1
; GENERAL INFORMATION:
; APPLICANT: Leite, Mario
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Fernandes, Elma
; APPLICANT: Li, Li
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Liu, Xiaohong
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Patturajan, Meera
; APPLICANT: Blalock, Angela
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Malvankar, Uriel M
; APPLICANT: Gusev, Vladimir
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter S
; APPLICANT: Ellerman, Karen
; APPLICANT: Heyes, Melvin P
; APPLICANT: Herrman, John
; APPLICANT: Pena, Carol E A
; APPLICANT: Shinkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Moore, No. US20040010119A11le
; APPLICANT: Shenoy, Suresh
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, Dave
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glenna
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-269
; CURRENT APPLICATION NUMBER: US/10/074,978A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: 60/268,221
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/335,109
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/312,284
; PRIOR FILING DATE: 2001-08-14

```

```

; PRIOR APPLICATION NUMBER: 60/268,496
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/276,703
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/330,293
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/322,127
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/280,899
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/310,797
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/268,646
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-074-978A-83

Query Match      100.0%; Score 20; DB 15; Length 2233;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACCGCATAGACTTCTCAGA 20
Db      1654 GACCGCATAGACTTCTCAGA 1673

RESULT 7
US-10-262-839-129
; Sequence 129, Application US/10262839
; Publication No. US20040038877A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John,
; APPLICANT: Anderson, David W.,
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Catterton, Elina,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Leach, Martin,
; APPLICANT: Li, Li,
; APPLICANT: Miller, Charles,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Shinkets, Richard,
; APPLICANT: Smithson, Glenna,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Vernet, Corine,
; APPLICANT: Voss, Edward,
; APPLICANT: Zerhusen, Brian,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462A
; CURRENT APPLICATION NUMBER: US/10/262,839
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09

```

```
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 129
; LENGTH: 2284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1768)
US-10-262-839-129

Query Match          100.0%; Score 20; DB 12; Length 2284;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20
    |||||
Db 1762 GACCGCATAGACTTCTCAGA 1781

RESULT 8
US-10-342-887-932
; Sequence 932, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 932
; LENGTH: 2326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-932

Query Match          100.0%; Score 20; DB 12; Length 2326;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20
    |||||
Db 1747 GACCGCATAGACTTCTCAGA 1766

RESULT 9
US-10-205-823-330
; Sequence 330, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Marjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wensey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 2326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-330

Query Match          100.0%; Score 20; DB 14; Length 2326;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20
    |||||
Db 1747 GACCGCATAGACTTCTCAGA 1766

RESULT 10
US-09-284-320-71
; Sequence 71, Application US/09284320
; Publication No. US20030092175A1
; GENERAL INFORMATION:
; APPLICANT: Kato, Seishi et al.
; TITLE OF INVENTION: HUMAN PROTEINS HAVING TRANSMEMBRANE DOMAINS AND DNAs
; TITLE OF INVENTION: ENCODING THESE PROTEINS
; FILE REFERENCE: GIN-6705CPUS
; CURRENT APPLICATION NUMBER: US/09/284,320
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: JP 8-301429
; PRIOR FILING DATE: 1996-11-13
; PRIOR APPLICATION NUMBER: PCT/JP97/04056
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 2373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

; LOCATION: (134)..(1810)  
US-09-284-320-71

Query Match 100.0%; Score 20; DB 10; Length 2373;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1804 GACCGCATAGACTTCTCAGA 1823

## RESULT 11

US-09-764-891-8086/c  
; Sequence 8086, Application US/09764891  
; Publication No. US20030077808A1

; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8086  
; LENGTH: 3160  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-764-891-8086

Query Match 100.0%; Score 20; DB 10; Length 3160;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1045 GACCGCATAGACTTCTCAGA 1026

## RESULT 12

US-09-764-891-8085/c  
; Sequence 8085, Application US/09764891  
; Publication No. US20030077808A1

; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8085  
; LENGTH: 7709  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-764-891-8085

Query Match 100.0%; Score 20; DB 10; Length 7709;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 1043 GACCGCATAGACTTCTCAGA 1024

## RESULT 13

US-09-972-546-15/c  
; Sequence 15, Application US/09972546  
; Publication No. US20030124704A1

; GENERAL INFORMATION:

; APPLICANT: STRITTMATTER, STEPHEN M.  
; APPLICANT: CAPE, RICHARD L.  
; APPLICANT: SAH, DINAH W.Y.  
; TITLE OF INVENTION: NOGO RECEPTOR HOMOLOGS  
; FILE REFERENCE: A116US  
; CURRENT APPLICATION NUMBER: US/09/972,546  
; CURRENT FILING DATE: 2001-10-06  
; PRIOR APPLICATION NUMBER: 60/238,361  
; PRIOR FILING DATE: 2000-10-06  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 143899  
; TYPE: DNA  
; ORGANISM: Homo sapiens

; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (2044)..(2144)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (6609)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (6625)..(6724)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (14153)..(14252)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (19512)..(19611)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (22595)..(22694)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (27825)..(27924)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (34953)..(35052)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (40783)..(40882)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (49000)..(49099)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (62884)..(62983)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (75528)..(75627)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (87944)..(88043)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (111030)..(111129)  
; OTHER INFORMATION: a, t, c, g, other or unknown

US-09-972-546-15

Query Match 100.0%; Score 20; DB 10; Length 143899;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 38586 GACCGCATAGACTTCTCAGA 38567

## RESULT 14

US-10-027-632-223788/c  
; Sequence 223788, Application US/10027632  
; Publication No. US20030204075A9

Search completed: April 10, 2004, 20:21:43  
Job time : 10.8778 secs

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 223788
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-223788
```

```
Query Match      82.0%; Score 16.4; DB 15; Length 642;
Best Local Similarity 94.4%; Pred. No. 81;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  GACCGCATAGACTTCTCA 18
DB      544 GACCATAGACTTCTCA 527
```

```
RESULT 15
US-09-728-446-1456/c
; Sequence 1456, Application US/09728446
; Patent No. US20020081668A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020081668A1 Murine Polynucleotide Sequences
;   AND MUTANT CELLS AND MUTANT ANIMALS DEFINED THEREBY
; FILE REFERENCE: LEX-0101-USA
; CURRENT APPLICATION NUMBER: US/09/728,446
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,270
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 1461
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1456
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(539)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-446-1456
```

```
Query Match      79.0%; Score 15.8; DB 9; Length 539;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  ACCGCATAGACTTCTCAGA 20
DB      329 ACCGCATAGACTTCTCTGA 311
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM' nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:16:47 ; Search time 68.1457 Seconds  
(without alignments)  
8764.220 Million cell updates/sec

Title: US-09-743-825-10  
Perfect score: 20  
Sequence: 1 gaccgcagatctctcaga 20

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hct:*
9: gb_est1:*
10: gb_est2:*
11: gb_hct:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pig:*
27: em_gss_vrl:*
28: gb_gse1:*
29: gb_gse2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	248	9	AA229970
C 2	20	100.0	302	10	AW136898
C 3	20	100.0	330	9	AA345572
C 4	20	100.0	345	9	AA448887

C	5	20	100.0	420	14	R00504
C 6	20	100.0	425	9	AA228376	nc37a02.s
C 7	20	100.0	457	9	AA004521	2489h01.s
C 8	20	100.0	468	14	H50956	Yp84c01.r1
C 9	20	100.0	512	9	AA228367	nc39d08.s
C 10	20	100.0	529	14	CF141328	UI-H-CB0
C 11	20	100.0	571	13	BX383685	BX383685
C 12	20	100.0	582	9	AA176965	nc03e06.s
C 13	20	100.0	586	12	BM725383	UI-E-EJ0-
C 14	20	100.0	613	12	BM677811	UI-E-EJ0-
C 15	20	100.0	617	10	AA956686	EST368756
C 16	20	100.0	656	12	BQ005970	UI-H-E11-
C 17	20	100.0	665	12	BQ006067	UI-H-E11-
C 18	20	100.0	677	13	BX103969	BX103969
C 19	20	100.0	726	10	B8872354	601448607
C 20	20	100.0	815	10	B8664173	602145978
C 21	20	100.0	823	10	BG032562	602301478
C 22	20	100.0	834	10	BF673273	602136120
C 23	20	100.0	837	12	BG497880	602543180
C 24	20	100.0	884	13	BU182024	AGENCOURT
C 25	20	100.0	923	10	BE293435	600942817
C 26	20	100.0	982	9	AL580746	AL580746
C 27	20	100.0	1171	13	BX447961	BX447961
C 28	20	100.0	1201	9	AL581753	AL581753
C 29	19.6	98.0	1201	9	AL583366	AL583366
C 30	19	95.0	169	9	AA040486	zf06e03.r
C 31	18.4	92.0	622	10	AW967482	EST379557
C 32	17.4	87.0	280	29	TA51811P	T. brucei
C 33	17	85.0	214	13	BQ301150	CM4-HB002
C 34	16.8	84.0	125	10	AW804677	QV4-UM009
C 35	16.8	84.0	341	9	AA933597	om55g03.s
C 36	16.8	84.0	507	10	AW139951	UI-H-B11-
C 37	16.8	84.0	533	14	CA378468	657359 NC
C 38	16.8	84.0	662	29	CE812330	tigr-gss-
C 39	16.8	84.0	673	12	BM727967	UI-E-EJ0-
C 40	16.8	84.0	733	29	CE833759	tigr-gss-
C 41	16.8	84.0	1005	28	CC264102	CH261-97B
C 42	16.8	84.0	1143	13	BX332990	BX332990
C 43	16.8	84.0	1392	29	AV400615	Pan trogl
C 44	16.4	82.0	370	28	A2825636	2M0100M22
C 45	16.4	82.0	421	9	AI117304	ub81f01.r

#### ALIGNMENTS

RESULT 1  
AA229970/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

nc37a02.s1 NCI\_CGAP\_Pr2 Homo sapiens cDNA clone IMAGE:1010282, mRNA  
248 bp mRNA linear EST 21-AUG-1997  
sequence.  
AA229970  
AA229970.1 GI:1852263  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 248)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapsb-remail.nih.gov](mailto:cgapsb-remail.nih.gov)  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuauqui,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)

Seq primer: -41m13 fwd. ET from Amersham  
High quality sequence stop: 215.

# FEATURES

Location/Qualifiers  
1..248  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1010282"  
/sex="Male"  
/dev\_stages="45 years old"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP\_Pr2"  
/vector="pAMP10; Site 1: Not1; Site 2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (life technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

# ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 248;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 180 GACCGCATAGACTTCTCAGA 161

# RESULT 2

AW136898/c  
LOCUS AW136898 302 bp mRNA linear EST 29-OCT-1999  
DEFINITION UI-H-B11-adp-b-02-0-UI.s1 NCI\_CGAP\_Sub3 Homo sapiens cDNA clone  
IMAGE:2717499 3', mRNA sequence.  
ACCESSION AW136898  
VERSION AW136898.1 GI:6141031  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 302)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: M13 Forward  
POLYA=Yes

# FEATURES

Location/Qualifiers  
1..302  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2717499"  
/lab\_host="DH10B (life Technologies)"  
/clone\_lib="NCI\_CGAP\_Sub3"  
/note="Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The

NCI CGAP Sub3 library is a subtracted library derived from the NCI\_CGAP Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalized or subtracted NCI\_CGAP libraries:  
NCI\_CGAP\_Co4, NCI\_CGAP\_Pr22, NCI\_CGAP\_Pr28, NCI\_CGAP\_Co10, NCI\_CGAP\_Co16, NCI\_CGAP\_Kid5, NCI\_CGAP\_Kid12, NCI\_CGAP\_Kid3, NCI\_CGAP\_Kid11, NCI\_CGAP\_Lym2, NCI\_CGAP\_Br2, NCI\_CGAP\_Co8, NCI\_CGAP\_CLL1, NCI\_CGAP\_Lei2, NCI\_CGAP\_Brn23, NCI\_CGAP\_Lu5, NCI\_CGAP\_Lu24, NCI\_CGAP\_Lu19, NCI\_CGAP\_GC4, NCI\_CGAP\_GC6, NCI\_CGAP\_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:  
NCI\_CGAP\_Kid3 pool 1 L1AM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clones 132376-132391, 145608-145677, 150052-150285); NCI\_CGAP\_Kid5 pool 1 L1AM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clones 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI\_CGAP\_Lu5 pool 1 L1AM 3575-3582, 3851-3854 (IMAGE Clones 1414920-1417991, 1520904-1522439); NCI\_CGAP\_GC4 pool 1 L1AM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clones 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI\_CGAP\_Pr22 pool 1 L1AM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clones 985608-986759, 1101192-1101959, 1217928-1220615); NCI\_CGAP\_Co10 pool 1 L1AM 2644-2653, 2871-2872 (IMAGE Clones 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described (Bonaldi, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.  
TAG TISSUE=colon  
TAG\_LIB=NCI\_CGAP\_Co10  
TAG\_SEQ=AAACG"

# ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 302;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 180 GACCGCATAGACTTCTCAGA 161

# RESULT 3

AA345572 330 bp mRNA linear EST 21-APR-1997  
DEFINITION EST51617 Gall bladder II Homo sapiens cDNA 5' end, mRNA sequence.  
ACCESSION AA345572  
VERSION AA345572.1 GI:1997820  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 330)

# REFERENCE

AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.P., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkley,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,X., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dinke,D., Feng,D.-F., Fertie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,



TITLE Fraser, C.M. and Venter, J.C.  
 JOURNAL Initial assessment of human gene diversity and expression patterns  
 MEDLINE based upon 83 million nucleotides of cDNA sequence  
 PUBMED Nature 377 (6547 Suppl), 3-174 (1995)  
 COMMENT 96026280

Other ESTs: THC148666  
 Contact: Kerlavage, AR  
 Bioinformatics  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850 USA  
 Tel: 3018699056  
 Fax: 3018699423  
 Email: arkerlav@tigr.org  
 For clone availability, additional sequence and expression  
 information related to this EST, please check the TIGR Human Gene  
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
 Seq primer: M13 Reverse.

#### FEATURES

source  
 1. .330  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="ATCC (inhost):147107"  
 /db\_xref="taxon:9606"  
 /sex="female"  
 /dev\_stage="adult, 25 yrs"  
 /clone\_lib="Gall bladder II"  
 /note="Organ: gall bladder; Vector: pBluescript SK-;  
 Site\_1: EcoRI; Site\_2: XhoI"

#### ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
 |||||  
 Db 300 GACCGCATAGACTTCTCAGA 319

#### RESULT 4

AA448887/c  
 LOCUS zxl1f09.s1 Soares total fetus Nb2HF8\_9w Homo sapiens cDNA clone  
 DEFINITION IMAGE:786185 3', mRNA sequence.

ACCESSION AA448887  
 VERSION AA448887.1 GI:2162557  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 345)

#### REFERENCE

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
 Kucaba, T., Lacy, M., Lennon, G., Marra, M., Martin, J.,  
 Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,  
 White, Y., Wylie, T., Waterston, R. and Wilson, R.

#### TITLE

JOURNAL The WashU-Merck EST Project 1997  
 COMMENT Unpublished (1997)  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.llnl.gov](http://info@image.llnl.gov)) for further information.

Seq primer: -41m13 fwd. ET from Amersham  
 High quality sequence stop: 299.

#### FEATURES

source  
 1. .345  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"

/db\_xref="GDB:5983804"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:786185"  
 /dev\_stage="8-9 weeks"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares total fetus Nb2HF8\_9w"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from pooled 8-9 week  
 (total) fetus material with a Not I - oligo(dT) primer [5'  
 TGTACCAATCTGAAGTGGAGCGCCGCTTAATTTTTTTTTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT7T3 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

#### ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 345;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
 |||||  
 Db 146 GACCGCATAGACTTCTCAGA 127

#### RESULT 5

R00504/c  
 LOCUS ye74c07.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
 DEFINITION IMAGE:123468 3', mRNA sequence.

ACCESSION R00504  
 VERSION R00504.1 GI:750240  
 KEYWORDS EST.

#### SOURCE

Homo sapiens (human)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 420)

#### REFERENCE

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
 Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and  
 Wilson, R.

#### TITLE

JOURNAL The WashU-Merck EST Project  
 COMMENT Unpublished (1995)  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert Size: 599

High quality sequence stops: 321 Source: IMAGE Consortium, LNL  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.llnl.gov](http://info@image.llnl.gov)) for further information.  
 Insert Length: 599 Std Error: 0.00  
 Seq primer: -21m13  
 High quality sequence stop: 321.

#### FEATURES

source  
 1. 420  
 Location/Qualifiers

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:476013"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:123468"  
 /sex="male"

/dev\_stage="20 week-post conception fetus"  
 /lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares fetal liver spleen INFLS"  
 /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)  
 with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;

1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACGGAGATTAAAGATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 420;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
|||||

Db 164 GACCGCATAGACTTCTCAGA 145  
|||||

## RESULT 6

AA228376/c  
LOCUS AA228376 425 bp mRNA linear EST 21-AUG-1997  
DEFINITION nc39e09.s1 NCI\_CGAP\_Pr2 Homo sapiens cDNA clone IMAGE:1010536, mRNA sequence.

ACCESSION AA228376

VERSION AA228376.1 GI:1849927

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 425)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

## JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,

M.D., Michael Emert-Buck, M.D., Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.

cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Seq primer: -4im3 fwd. Et from Amersham

High quality sequence stop: 403.

## FEATURES

Location/Qualifiers

1..425

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1010536"

/sex="Male"

/dev\_stage="45 years old"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Pr2"

/note="Vector: pAMP10; Site 1: NotI; Site 2: EcoRI; 1st

strand cDNA was primed with oligo(dT)17 on 50 ng of

DNase-treated, total cellular RNA obtained from

5,000-10,000 microdissected preneoplastic cells

histologically-determined to be prostatic intraepithelial

neoplasia 2 (PIN2) cells. Double-stranded cDNA was

ligated to EcoRI adaptors, 5 cycles of PCR applied to the

cDNA with an adaptor-specific primer, and the resulting

PCR product subcloned into pAMP10 by the UDG-cloning

method (Life Technologies). Average insert size is 600

bp. NOTE: Not directionally cloned. This library was

constructed by David Krizman."

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 425;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
|||||

Db 180 GACCGCATAGACTTCTCAGA 161  
|||||

## RESULT 7

AA004521/c  
LOCUS AA004521 457 bp mRNA linear EST 07-MAY-1997  
DEFINITION zh98h01.s1 Soares\_fetal\_liver\_spleen\_lnfpls\_s1 Homo sapiens cDNA clone IMAGE:428497 3', mRNA sequence.

ACCESSION AA004521

VERSION AA004521.1 GI:1448098

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 457)

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,

Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,

Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,

Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,

Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.

and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478

PUBMED 889549

## COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

Insert Length: 1202 Std Error: 0.00

Seq primer: mob.REGA+ET.

## FEATURES

Location/Qualifiers

1..457

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:1328266"

/db\_xref="taxon:9606"

/clone="IMAGE:428497"

/sex="male"

/dev\_stage="20 week-post conception fetus"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares\_fetal\_liver\_spleen\_lnfpls\_s1"

/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)

with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;

This is a subtracted version of the original Soares fetal

liver spleen lnfpls library. 1st strand cDNA was primed

with a Pac I - oligo(dT) primer [5'

AACGGAGATTAAAGATCTTTTCTTTTCTTTT 3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 457;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACCGCATAGACTTCTCAGA 20  
|||||

Db 144 GACCGCATAGACTTCTCAGA 125  
|||||



PUBMED  
COMMENT

8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Tim Ratliff  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/humanfl.html>  
Seq primer: pYX-5.

FEATURES  
source

Location/Qualifiers  
1..529  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3098904"  
/issue\_type="CNCAP(3)T-225 cell line"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH MGC\_210"  
/note="Organ: Prostate; Vector: p7T73 Pac; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with Not I and then cloned  
directionally into p7T73 Pac vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CCCAC. Tissue was provided by Tim Ratliff."

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 529;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 97 GACCGCATAGACTTCTCAGA 116

RESULT 11  
BX383685

LOCUS BX383685 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED  
DEFINITION Homo sapiens cDNA clone CSODJ014YD24 5-PRIME, mRNA sequence.

ACCESSION BX383685  
VERSION BX383685.1 GI:30444078

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)

## COMMENT

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6149.r For  
more information about this cluster, see  
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODJ014DB120P1&cluster=6149.r)  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODJ014DB120P1&cluster=6149.r>. Contact :  
Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

FEATURES  
source

Faraday Avenue Genoscope sequence ID : CSODJ014DB12QP1.  
Location/Qualifiers  
1..571  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODJ014YD24"  
/cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
/cell\_line="JURKAT"  
/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT  
10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 571;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 441 GACCGCATAGACTTCTCAGA 460

## RESULT 12

AA176965/c

LOCUS AA176965

DEFINITION NC030606.81 NCI\_CGAP\_Pr2 Homo sapiens cDNA clone IMAGE:342, mRNA  
sequence.

ACCESSION AA176965

VERSION AA176965.1 GI:1758188

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL Unpublished (1997)

## COMMENT

Contact: Robert Straubeberg, Ph.D.  
Email: [csapbs@mail.nih.gov](mailto:csapbs@mail.nih.gov)  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 203.

## FEATURES

## source

Location/Qualifiers  
1..582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:342"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Pr2"  
/note="Vector: pAMP10; Site 1: NotI; Site 2: EcoRI; 1st  
strand cDNA was primed with oligo(dT)17 on 50 ng of  
DNase-treated, total cellular RNA obtained from  
5,000-10,000 microdissected preneoplastic cells  
histologically-determined to be prostatic intraepithelial  
neoplasia 2 (PIN2) cells. Double-stranded cDNA was  
ligated to EcoRI adaptors, 5 cycles of PCR applied to the  
cDNA with an adaptor-specific primer, and the resulting

PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 582;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
DB 180 GACCGCATAGACTTCTCAGA 161

## RESULT 13

## LOCUS

BM725383 586 bp mRNA linear EST 01-MAR-2002  
UI-E-EJO-aie-1-15-0-UI.r1 UI-E-EJO Homo sapiens cDNA clone  
UI-E-EJO-aie-1-15-0-UI 5', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Reverse.  
Location/Qualifiers  
1. .586  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJO-aie-1-15-0-UI"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"

/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-EJO"  
/notes="Organ: eye; Vector: p7T73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EJO is a subcloned cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AATGCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 586;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
DB 14 GACCGCATAGACTTCTCAGA 33

## RESULT 14

## LOCUS

BM677811 613 bp mRNA linear EST 27-FEB-2002  
UI-E-EJO-aie-1-15-0-UI.r1 UI-E-EJO Homo sapiens cDNA clone  
UI-E-EJO-aie-1-15-0-UI 3', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Forward  
Location/Qualifiers  
1. .613  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJO-aie-1-15-0-UI"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"

/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-EJO"  
/notes="Organ: eye; Vector: p7T73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EJO is a subcloned cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

located between the Not 1 site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGATCAGA; lens, CGATTAGCA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTG; retina, CCCTG, Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). TAG TISSUE-RPE and Choroid TAG\_LIB=UI-E-EJ0 TAG\_SEQ=ACCTA"

# ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 613;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 591 GACCGCATAGACTTCTCAGA 572

## RESULT 15

AW956686  
LOCUS EST368756 617 bp mRNA linear EST 01-JUN-2000  
DEFINITION MAGD mouse cDNA, mRNA sequence.  
ACCESSION AW956686  
VERSION AW956686.1 GI:8146369  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 617)

AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,  
Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and  
Quackenbush, J.

TITLE Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray

JOURNAL Unpublished (2000)  
COMMENT Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnq@tigr.org  
Plate: 95

Seq primer: Reverse.

## FEATURES

source  
Location/Qualifiers  
1..617  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGD"  
/note="Vector: pBluescriptSKm"

# ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 617;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCGCATAGACTTCTCAGA 20  
|||||  
Db 299 GACCGCATAGACTTCTCAGA 318

Search completed: April 10, 2004, 19:48:55  
Job time : 69.3957 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 14:24:52 ; Search time 9.22729 seconds  
(without alignments)  
10128.688 Million cell updates/sec

Title: US-09-743-825-7

Perfect score: 22

Sequence: 1 gcattacagtagaagcc 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001s:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	3 AAZ50444	Aaz50444 EST R0050
2	22	100.0	457	6 ABN93510	Abn93510 Gene #8 u
3	22	100.0	856	4 AAL01556	Aal01556 Human rep
4	22	100.0	856	4 ABL96993	AbL96993 Human tes
5	22	100.0	1737	4 AAL05399	Aal05399 Human rep
6	22	100.0	1737	4 ABL98261	AbL98261 Human tes
7	22	100.0	2093	4 AAH16350	Aah16350 Human cDN
8	22	100.0	2105	7 ACC72138	Acc72138 Human NOV
9	22	100.0	2233	7 ACD20413	Acd20413 DNA encod
10	22	100.0	2284	7 ACC72139	Acc72139 Human NOV
11	22	100.0	2326	3 AAZ50442	Aaz50442 Human PB3
12	22	100.0	2326	9 ADB75506	AdB75506 Prostate
13	22	100.0	2373	2 AAU49590	Aau49590 Human liv
14	22	100.0	3160	4 AAL05398	Aal05398 Human rep
15	22	100.0	3160	4 ABL98260	AbL98260 Human tes
16	22	100.0	7709	4 AAL05397	Aal05397 Human rep
17	22	100.0	7709	4 ABL98259	AbL98259 Human tes
18	22	100.0	143899	6 AAL138336	Aal138336 Genomic s
19	18.8	85.5	579	4 AAH12098	Aah12098 Human cDN
20	18	81.8	2355	7 ACF71472	Acf71472 Phototrab
21	18	81.8	110000	7 ACF67367	Continuation (47 o
22	18	81.8	110000	7 ACF65388	Continuation (2 of
23	17.8	80.9	6454	4 ABL11966	AbL11966 Drosophil

## ALIGNMENTS

### RESULT 1

AAZ50444

ID AAZ50444 standard; DNA; 22 BP.

XX

AC AAZ50444;

DT 18-MAY-2000 (first entry)

XX

DE EST R00504-specific primer. 1.

XX

PB39; human; prostate cancer; PC; chromosome 11p11.1-11.2; cancer;  
prostate epithelium; splicing mechanism; early diagnosis; progression;  
precancerous cell; metastatic potential; non-neoplastic prostate disease;  
expressed sequence tag; EST; PCR primer; ss.

OS Homo sapiens.

XX

XX WO200005376-A1.

XX

PD 03-FEB-2000.

XX

PF 23-JUL-1999; 99WO-US016831.

XX

PR 24-JUL-1998; 98US-0094137P.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Chuasqui RF, Cole KA, Liotta LA;

XX

DR WPI; 2000-182700/16.

XX

Novel gene which is dysregulated in prostate cancer useful for diagnosing cancer.

PT

XX

PS Claim 5; Page 16; 51pp; English.

XX

The present sequence is the EST AAR00504-specific PCR primer, used for amplification of sequences contained within the EST AAR00504. It is useful to probe the gene overexpressed in prostate cancer epithelium and to analyse the differential expression of the EST. The PB39 gene that is dysregulated in prostate cancer is isolated from human pancreas cDNA library and has homology to the EST AAR00504. PB39 gene is located on chromosome 11p11.1-11.2. Abnormally high concentrations of PB39 are found in prostate tissue derived from prostate cancer (PC) epithelium. PB39 sequence is useful for detection of precancerous or cancer cells in the prostate. PB39 is useful for early diagnosis of the progression of

AbL17168 Drosophil  
Abt03288 Ovary cel  
Abt03289 Ovary cel  
Abk34854 Human CDN  
Adb62134 Human CDN  
Aca47120 Prokaryot  
Abn92775 Staphyloc  
Abk14717 cDNA enco  
AbL19878 Drosophil  
Aah54059 S. epider  
Aak78529 Human imm  
Abk69947 Human sec  
Abz68144 Human sec  
Abz74623 Secreted  
Adc21014 Human sec  
Aak78528 Human imm  
Abk69948 Human sec  
Abz68145 Human sec  
Abz74624 Secreted  
Adc21015 Human sec  
AbL04034 Drosophil

C 24 17.8 80.9 9489 4 ABL17168  
25 17.4 79.1 266 6 ABT03288  
26 17.4 79.1 1483 6 ABT03289  
C 27 17.4 79.1 1630 6 ABK34854  
C 28 17.4 79.1 2005 9 ADB62134  
29 17.2 78.2 2939 9 ADB68931  
30 16.8 76.4 1089 7 ACA47120  
31 16.8 76.4 1098 6 ABN92775  
C 32 16.8 76.4 2895 6 ABK14717  
33 16.8 76.4 3493 4 ABL19878  
C 34 16.8 76.4 3506 4 AAH54059  
C 35 16.8 76.4 16037 6 AAK78529  
C 36 16.8 76.4 16037 6 ABK69947  
C 37 16.8 76.4 16037 7 ABZ68144  
C 38 16.8 76.4 16037 7 ABZ74623  
C 39 16.8 76.4 16037 9 ADC21014  
C 40 16.8 76.4 16107 6 AAK78528  
C 41 16.8 76.4 16107 6 ABK69948  
C 42 16.8 76.4 16107 7 ABZ68145  
C 43 16.8 76.4 16107 7 ABZ74624  
C 44 16.8 76.4 16107 9 ADC21015  
45 16.8 76.4 17478 4 ABL04034

CC prostate cancer, especially in aggressive prostate carcinoma. It can also  
 CC distinguish PC from other non-neoplastic prostate disease. The diagnostic  
 CC method is selective and specific for various types of PC and also  
 CC facilitates identifying prostate cancer of differing aggressiveness and  
 CC metastatic potential  
 XX  
 SQ Sequence 22 BP; 8 A; 4 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 3; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 0.37;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCATGTTACAGGTAGAAAAGCC 22  
 |||||  
 Db 1 GCATGTTACAGGTAGAAAAGCC 22

RESULT 2  
 ABN93510  
 ID ABN93510 standard; DNA; 457 BP.  
 XX  
 AC ABN93510;  
 XX  
 DT 13-AUG-2002 (first entry)  
 XX  
 DE Gene #8 used to diagnose liver cancer.  
 XX  
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
 KW metastatic liver tumour; cytostatic; expression profile; disease state;  
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200229103-A2.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 02-OCT-2001; 2001WO-US030599.  
 XX  
 PR 02-OCT-2000; 2000US-0237054P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 XX  
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
 XX  
 DR WPI; 2002-426119/45.  
 XX

Diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, involves detecting the level of expression of two or more genes in a liver tissue sample.  
 Claim 1; SEQ ID NO 8; 298pp; English.  
 The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN93503-ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytostatic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Query Match 100.0%; Score 22; DB 6; Length 457;  
 Best Local Similarity 100.0%; Pred. No. 0.56;

Sequence 457 BP; 117 A; 108 C; 119 G; 113 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 22; DB 6; Length 457;  
 Best Local Similarity 100.0%; Pred. No. 0.56;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GCATGTTACAGGTAGAAAAGCC 22  
 |||||  
 Db 44 GCATGTTACAGGTAGAAAAGCC 65  
 RESULT 3  
 AAL01556  
 ID AAL01556 standard; cDNA; 856 BP.  
 XX  
 AC AAL01556;  
 XX  
 DT 21-NOV-2001 (first entry)  
 XX  
 DE Human reproductive system related antigen cDNA SEQ ID NO: 1557.  
 XX  
 KW Human; reproductive system related antigen; reproductive system disorder;  
 KW cancer; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155320-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001339.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205151P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226686P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.



PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0246417P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465570/50.  
DR P-PSDB; AAM95586.  
XX  
PT Isolated nucleic acid molecule encoding a reproductive system antigen is  
PT used in preventing, treating or ameliorating a medical condition.  
XX  
PS Claim 1; SEQ ID NO 1557; 1297pp + Sequence Listing; English.  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a coding sequence of the  
CC invention  
XX  
SQ Sequence 856 BP; 209 A; 208 C; 234 G; 203 T; 0 U; 2 Other;  
Query Match 100.0%; Score 22; DB 4; Length 856;  
Best Local Similarity 100.0%; Pred. No. 0.61; Mismatches 0; Gaps 0;  
Matches 22; Conservative 0; Indels 0; Indels 0; Gaps 0;  
QY 1 GCATGTTACAGGTAGAAAGCC 22  
DB 63 GCATGTTACAGGTAGAAAGCC 84  
RESULT 4  
ABL96993  
ID ABL96993 standard; cDNA; 856 BP.  
XX  
XX ABL96993;  
XX  
XX 21-JUN-2002 (first entry)  
XX Human testicular antigen encoding cDNA SEQ ID NO: 661.  
XX  
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
XX reproductive system disorder; urinary system disorder; gene therapy;  
XX cardiovascular disorder; respiratory disorder; neurological disorder;  
XX gastrointestinal disease; infection; cytostatic; gene; ss.  
OS Homo sapiens.  
XX  
XX WO20015317-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001329.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0198874P.  
 PR 17-MAR-2000; 2000US-0190078P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216664P.  
 PR 07-JUL-2000; 2000US-0216680P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225265P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227003P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251858P.  
 PR 08-DEC-2000; 2000US-0251859P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-02559678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-483232/52.  
 XX  
 XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
 PT for preventing, diagnosing and/or treating testicular cancer.  
 XX  
 PS Claim 1; SEQ ID NO 661; 766pp; English.  
 XX

CC The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a cDNA of the  
CC invention

XX SQ Sequence 856 BP; 209 A; 208 C; 234 G; 203 T; 0 U; 2 Other;

Query Match 100.0%; Score 22; DB 4; Length 856;  
Best Local Similarity 100.0%; Pred. No. 0.61;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAGCC 22  
|||||  
Db 63 GCATGTTACAGGTAGAAAGCC 84

RESULT 5

AAL05399

ID AAL05399 standard; DNA; 1737 BP.

XX AAL05399;

XX 21-NOV-2001 (first entry)

XX Human reproductive system related antigen DNA SEQ ID NO: 8087.

KW Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ds.

XX Homo sapiens.

XX WO200155320-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001339.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.

```
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0253678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
PT
XX Disclosure; SEQ ID NO 8087; 1297pp + Sequence Listing; English.
XX
PS The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX
SQ Sequence 1737 BP; 443 A; 431 C; 458 G; 405 T; 0 U; 0 Other;
Query Match 100.08; Score 22; DB 4; Length 1737;
Best Local Similarity 100.08; Pred. No. 0.68;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCATGTTACAGGTAGAAAGCC 22
Db 944 GCATGTTACAGGTAGAAAGCC 965
RESULT 6
ABL98261
ID ABL98261 standard; DNA; 1737 BP.
XX
AC ABL98261;
XX
XX 21-JUN-2002 (first entry)
XX
XX Human testicular antigen encoding DNA fragment SEQ ID NO: 2913.
XX
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
```

```
XX
OS Homo sapiens.
XX
XX WO200155317-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001329.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 12-SEP-2000; 2000US-0232081P.
XX 14-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 21-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 25-SEP-2000; 2000US-0234998P.
```

26-SEP-2000; 2000US-0235484P.  
27-SEP-2000; 2000US-0235834P.  
27-SEP-2000; 2000US-0235836P.  
28-SEP-2000; 2000US-0236327P.  
29-SEP-2000; 2000US-0236367P.  
29-SEP-2000; 2000US-0236368P.  
29-SEP-2000; 2000US-0236369P.  
29-SEP-2000; 2000US-0236370P.  
02-OCT-2000; 2000US-0236802P.  
02-OCT-2000; 2000US-0237037P.  
02-OCT-2000; 2000US-0237038P.  
02-OCT-2000; 2000US-0237039P.  
02-OCT-2000; 2000US-0237040P.  
13-OCT-2000; 2000US-0239935P.  
13-OCT-2000; 2000US-0239937P.  
20-OCT-2000; 2000US-0240960P.  
20-OCT-2000; 2000US-0241221P.  
20-OCT-2000; 2000US-0241785P.  
20-OCT-2000; 2000US-0241786P.  
20-OCT-2000; 2000US-0241787P.  
20-OCT-2000; 2000US-0241808P.  
20-OCT-2000; 2000US-0241809P.  
20-OCT-2000; 2000US-0241826P.  
01-NOV-2000; 2000US-0244617P.  
08-NOV-2000; 2000US-0244647P.  
08-NOV-2000; 2000US-0246475P.  
08-NOV-2000; 2000US-0246476P.  
08-NOV-2000; 2000US-0246477P.  
08-NOV-2000; 2000US-0246478P.  
08-NOV-2000; 2000US-0246523P.  
08-NOV-2000; 2000US-0246532P.  
08-NOV-2000; 2000US-0246533P.  
08-NOV-2000; 2000US-0246532P.  
08-NOV-2000; 2000US-0246610P.  
08-NOV-2000; 2000US-0246611P.  
08-NOV-2000; 2000US-0246613P.  
17-NOV-2000; 2000US-0249207P.  
17-NOV-2000; 2000US-0249208P.  
17-NOV-2000; 2000US-0249209P.  
17-NOV-2000; 2000US-0249210P.  
17-NOV-2000; 2000US-0249211P.  
17-NOV-2000; 2000US-0249212P.  
17-NOV-2000; 2000US-0249213P.  
17-NOV-2000; 2000US-0249214P.  
17-NOV-2000; 2000US-0249215P.  
17-NOV-2000; 2000US-0249216P.  
17-NOV-2000; 2000US-0249217P.  
17-NOV-2000; 2000US-0249218P.  
17-NOV-2000; 2000US-0249244P.  
17-NOV-2000; 2000US-0249245P.  
17-NOV-2000; 2000US-0249264P.  
17-NOV-2000; 2000US-0249265P.  
17-NOV-2000; 2000US-0249297P.  
17-NOV-2000; 2000US-0249299P.  
17-NOV-2000; 2000US-0249300P.  
01-DEC-2000; 2000US-0250160P.  
01-DEC-2000; 2000US-0250391P.  
05-DEC-2000; 2000US-0251030P.  
05-DEC-2000; 2000US-0251988P.  
06-DEC-2000; 2000US-0256719P.  
06-DEC-2000; 2000US-0251479P.  
08-DEC-2000; 2000US-0251856P.  
08-DEC-2000; 2000US-0251868P.  
08-DEC-2000; 2000US-0251869P.  
08-DEC-2000; 2000US-0251989P.  
11-DEC-2000; 2000US-0251990P.  
05-JAN-2001; 2001US-0259678P.  
XX

PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI; 2001-483232/52.  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
PT for preventing, diagnosing and/or treating testicular cancer.  
XX  
XX Disclosure; SEQ ID NO 2913; 766pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention  
XX  
SQ Sequence 1737 BP; 443 A; 431 C; 458 G; 405 T; 0 U; 0 Other;  
Query Match 100.0%; Score 22; DB 4; Length 1737;  
Best Local Similarity 100.0%; Pred. No. 0.68; Mismatches 0; Indels 0; Gaps 0;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
DB 944 GCATGTTACAGGTAGAAAAGCC 965  
RESULT 7  
AAH16350/c  
ID AAH16350 standard; cDNA; 2093 BP.  
XX  
AC AAH16350;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human cDNA sequence SEQ ID NO:15276.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX Homo sapiens.  
XX  
PN BP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
XX 28-JUL-2000; 2000EP-00116126.  
XX  
XX 29-JUL-1999; 99JP-00248036.  
XX 27-AUG-1999; 99JP-00300253.  
XX 11-JAN-2000; 2000JP-00118776.  
XX 02-MAY-2000; 2000JP-00183767.  
XX 09-JUN-2000; 2000JP-00241899.  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
XX WPI; 2001-318749/34.  
XX  
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
PT length cDNAs defined in the specification, and for the detection and/or  
PT diagnosis of the abnormality of the proteins encoded by the full-length  
PT cDNAs.  
XX  
XX Claim 8; SEQ ID NO 15276; 2537pp + Sequence Listing; English.  
XX  
XX The present invention describes primer sets for synthesising 5602 full-  
CC length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the

complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

Sequence 2093 BP; 429 A; 623 C; 548 G; 493 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 4; Length 2093;  
 Best Local Similarity 100.0%; Pred. No. 0.7;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCATGTTACAGGTAGAAAAGCC 22  
 |||||  
 Db 1625 GCATGTTACAGGTAGAAAAGCC 1604

RESULT 8  
 ACC72138/c  
 ID ACC72138 standard; DNA; 2105 BP.  
 XX  
 AC ACC72138;  
 XX  
 DT 07-JUL-2003 (first entry)  
 XX  
 DE Human NOV33a coding sequence.  
 XX  
 KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; metabolic disorder;  
 KW diabetes; obesity; infection; cachexia; cancer; dyslipidaemia;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; gene; ds.

Homo sapiens.  
 OS  
 PN WO2003029423-A2.  
 XX  
 XX 10-APR-2003.  
 PD  
 XX  
 PF 02-OCT-2002; 2002WO-US031358.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327342P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328499P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 12-APR-2002; 2002US-0371972P.  
 PR 12-APR-2002; 2002US-0371980P.  
 PR 17-APR-2002; 2002US-0373261P.

19-APR-2002; 2002US-0373805P.  
 23-APR-2002; 2002US-0374738P.  
 PR 16-MAY-2002; 2002US-0381101P.  
 PR 17-MAY-2002; 2002US-0381635P.  
 PR 29-MAY-2002; 2002US-038330P.  
 PR 01-OCT-2002; 2002US-00262839.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 XX Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;  
 PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
 PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;  
 PI Rothenberg ME, Shmukler RA, Smithson G, Spytek KA, Taupier RJ;  
 PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;  
 DR WPI; 2003-381625/36.  
 DR P-PSDB; ABR58426.  
 XX  
 PT NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
 PT dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 20; Page 209-210; 487pp; English.  
 XX  
 CC The present invention relates to novel human NOV proteins and their  
 CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
 CC proteins are useful in manufacturing a medicament for treating a syndrome  
 CC associated with a human disease. The NOV proteins and coding sequences  
 CC may be used to diagnose, treat or prevent metabolic disorders such as  
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
 CC disorders, haematopoietic disorders and various dyslipidaemias  
 CC  
 SQ Sequence 2105 BP; 425 A; 629 C; 556 G; 495 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 7; Length 2105;  
 Best Local Similarity 100.0%; Pred. No. 0.7;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCATGTTACAGGTAGAAAAGCC 22  
 |||||  
 Db 1682 GCATGTTACAGGTAGAAAAGCC 1661

RESULT 9  
 ACD20413/c  
 ID ACD20413 standard; DNA; 2233 BP.  
 XX  
 AC ACD20413;  
 XX  
 DT 26-AUG-2003 (first entry)  
 XX  
 DE DNA encoding human NOV28 protein.  
 XX  
 KW Human; NOVX; inflammatory disorder; demyelination disease; stroke;  
 KW renal disorder; infection; cardiomyopathy; atherosclerosis; acne;  
 KW hypercension; pancreatitis; Von Hippel-Lindau; endometriosis; fertility;  
 KW scleroderma; cirrhosis; inflammatory bowel disease; Crohn's disease;  
 KW haemophilia; autoimmune disease; allergy; AIDS;  
 KW graft versus host disease; Alzheimer's disease; arthritis; pain;  
 KW Parkinson's disease; Huntington's disease; obesity; diabetes;  
 KW hair growth; hair loss; asthma; schizophrenia; glomerulonephritis;  
 KW lupus erythematosus; psoriasis; antidiabetic; anorectic; metabolic;  
 KW nootropic; neuroprotective; cytostatic; antibacterial; virucide;  
 KW protozoicide; antiarteriosclerotic; hypotensive; cerebroprotective;  
 KW antiinflammatory; gynaecological; antinfertility; dermatological;  
 KW hepatotropic; haemostatic; immunosuppressive; antiallergic;  
 KW antiarthritic; anticonvulsant; antiseborrheic; antiasthmatic;  
 KW neuroleptic; anti-HIV; analgesic; nephrotropic; antipsoriatic; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX

PN WO200298917-A2.  
 XX 12-DEC-2002.  
 XX 12-FEB-2002; 2002WO-US022049.  
 XX 12-FEB-2001; 2001US-0268221P.  
 PR 13-FEB-2001; 2001US-0268496P.  
 PR 14-FEB-2001; 2001US-0268646P.  
 PR 14-FEB-2001; 2001US-0268655P.  
 PR 15-FEB-2001; 2001US-0269136P.  
 PR 16-FEB-2001; 2001US-0269310P.  
 PR 16-FEB-2001; 2001US-0269530P.  
 PR 15-MAR-2001; 2001US-0276405P.  
 PR 16-MAR-2001; 2001US-0276399P.  
 PR 16-MAR-2001; 2001US-0276703P.  
 PR 23-MAR-2001; 2001US-0278199P.  
 PR 28-MAR-2001; 2001US-0279274P.  
 PR 30-MAR-2001; 2001US-0280238P.  
 PR 02-APR-2001; 2001US-0280899P.  
 PR 08-AUG-2001; 2001US-0310797P.  
 PR 14-AUG-2001; 2001US-0312284P.  
 PR 14-SEP-2001; 2001US-0322294P.  
 PR 14-SEP-2001; 2001US-0322295P.  
 PR 18-OCT-2001; 2001US-0330293P.  
 PR 31-OCT-2001; 2001US-0335104P.  
 PR 31-OCT-2001; 2001US-0335109P.  
 PR 21-NOV-2001; 2001US-0332127P.  
 PR 28-NOV-2001; 2001US-0331772P.  
 XX  
 (CURA-) CURAGEN CORP.  
 XX Guo X, Fernandes E, Li L, Kekuda R, Liu Y, Leite M, Spytek KA;  
 PI Ji W, Casman SJ, Boldog FL, Patturajan M, Vernet CM, Ballinger RA;  
 PI Malyankar UM, Tchernov VT, Bialock AB, Gusev VY, Rastelli L;  
 PI Mezes PD, Ellerman K, Heyes M, Herrmann UL, Shinkets RA, Ioime N;  
 PI Pena CEA, Shenoy SG, Taupier RJ, Gerlach V, Gorman L;  
 XX WPI; 2003-148650/14.  
 DR P-PSDB; ABO15025.  
 XX  
 Novel NOVX polypeptide useful for identifying an agent that binds to the  
 PT polypeptide, and for treating cardiomyopathy, atherosclerosis,  
 PT hypertension, infertility, scleroderma, cirrhosis, and inflammatory bowel  
 PT disease.  
 XX  
 PS Claim 9; Page 217-218; 566pp; English.  
 XX  
 CC The present invention relates to the isolation of novel human  
 CC polypeptides referred to as NOVX (NOVI-NOV37), variants of these  
 CC proteins, and the polynucleotide sequences encoding them. The NOVX  
 CC proteins of the invention share homology to various types of protein  
 CC families such as zinc finger-like proteins, enzymes, receptors, and  
 CC lipoproteins. The sequences of the invention may be useful in the  
 CC manufacture of a medicament for treating a syndrome associated with a  
 CC human disease. For example they can be used to treat inflammatory  
 CC disorders, demyelination disease, renal disorders, infections,  
 CC cardiomyopathy, atherosclerosis, hypertension, stroke, pancreatitis, Von  
 CC Hippel-Lindau, endometriosis, fertility, scleroderma, cirrhosis,  
 CC inflammatory bowel disease, Crohn's disease, haemophilia, autoimmune  
 CC diseases, allergies, graft versus host disease, Alzheimer's disease,  
 CC arthritis, Parkinson's disease, Huntington's disease, obesity, diabetes,  
 CC acne, hair growth/loss, asthma, schizophrenia, AIDS, pain.  
 CC glomerulonephritis, lupus erythematosus, and psoriasis. ACD20372-ACD20427  
 CC represent DNA sequences encoding the NOVX polypeptides of the invention.  
 CC Note: SEQ ID Nos 113-460 are known sequences used for homology purposes  
 XX  
 SQ Sequence 2233 BP; 458 A; 659 C; 605 G; 511 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 22; DB 7; Length 2233;  
 Best Local Similarity 100.0%; Pred. No. 0.7;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAGCC 22  
 DB 1754 GCATGTTACAGGTAGAAAGCC 1733  
 RESULT 10  
 ACC72139/c  
 ID ACC72139 standard; DNA; 2284 BP.  
 XX  
 AC ACC72139;  
 XX 07-JUL-2003 (first entry)  
 XX Human NOV33b coding sequence.  
 XX Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; metabolic disorder;  
 KW diabetes; obesity; infection; cachexia; cancer; dyslipidemia;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; gene; ds.  
 XX Homo sapiens.  
 OS WO2003029423-A2.  
 XX 10-APR-2003.  
 XX 02-OCT-2002; 2002WO-US031358.  
 XX 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327342P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 12-APR-2002; 2002US-0371972P.  
 PR 12-APR-2002; 2002US-0371980P.  
 PR 17-APR-2002; 2002US-0373261P.  
 PR 19-APR-2002; 2002US-0373805P.  
 PR 23-APR-2002; 2002US-0374738P.  
 PR 16-MAY-2002; 2002US-0381101P.  
 PR 17-MAY-2002; 2002US-0381635P.  
 PR 29-MAY-2002; 2002US-0383830P.  
 PR 01-OCT-2002; 2002US-00262839.  
 XX (CURA-) CURAGEN CORP.  
 XX Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;  
 PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W; DK;  
 PI Kekuda R, Leach MD, Li L, Miller CB, Patturajan M, Rieger DJ;  
 PI Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ;  
 PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;  
 XX WPI; 2003-381625/36.  
 DR P-PSDB; ABR58427.  
 XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
 PT dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 20; Page 210-211; 487pp; English.  
 XX The present invention relates to novel human NOV proteins and their

CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
 CC proteins are useful in manufacturing a medicament for treating a syndrome  
 CC associated with a human disease. The NOV proteins and coding sequences  
 CC may be used to diagnose, treat or prevent metabolic disorders such as  
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
 CC disorders, haematopoietic disorders and various dyslipidaemias  
 XX  
 SQ Sequence 2284 BP; 446 A; 697 C; 605 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 7; Length 2284;  
 Best Local Similarity 100.0%; Pred. No. 0.7;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCATGTTACAGGTAGAAAAGCC 22

Db 1862 GCATGTTACAGGTAGAAAAGCC 1841

RESULT 11

AAZ50442/c

ID AAZ50442 standard; cDNA; 2326 BP.

XX AC AAZ50442;

XX DT 18-MAY-2000 (first entry)

XX DE Human PB39 cDNA, a gene dysregulated in prostate cancer.

XX KW PB39; human; prostate cancer; PC; chromosome 11p11.1-11.2; cancer;  
 KW prostate epithelium; splicing mechanism; early diagnosis; progression;  
 KW precancerous cell; aggressive prostate carcinoma; metastatic potential;  
 KW non-neoplastic prostate disease; expressed sequence tag; EST; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 XX 5'UTR 1..76  
 XX /tag= a  
 XX misc\_signal 71..80  
 XX /tag= b  
 XX /label= Kozak\_consensus\_sequence  
 XX CDS 77..1756  
 XX /tag= c  
 XX /product= "Human PB39 protein"  
 XX /note= "Over-expressed in prostate cancer"  
 XX 3'UTR 1757..2326  
 XX /tag= d

XX WO200005376-A1.

XX PD 03-FEB-2000.

XX PF 23-JUL-1999; 99WO-US016831.

XX PR 24-JUL-1998; 98US-0094137P.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Chuaqui RF, Cole KA, Liotta LA;

XX DR WPI; 2000-182700/16.

XX P-PSDB; AAY44897.

XX PT Novel gene which is dysregulated in prostate cancer useful for diagnosing  
 PT cancer.

XX PS Claim 1; Page 37-40; 51pp; English.

XX CC The present sequence is the human PB39 cDNA, a gene that is dysregulated  
 CC in prostate cancer and is isolated from human pancreas cDNA library. It  
 CC has homology to an expressed sequence tag (EST) AAR00504. It is mapped to  
 CC chromosome 11p11.1-11.2. A variant of PB39 results from an alternative

CC RNA splicing mechanism, yielding a larger transcript (5Kb). Abnormally  
 CC high concentrations of PB39 are found in prostate tissue derived from  
 CC prostate cancer (PC) epithelium. PB39 sequence is useful for detection of  
 CC precancerous or cancer cells in the prostate. PB39 is useful for early  
 CC diagnosis of the progression of prostate cancer, especially in aggressive  
 CC prostate carcinoma. It can also distinguish PC from other non-neoplastic  
 CC prostate disease. The diagnostic method is selective and specific for  
 CC various types of PC and also facilitates identifying prostate cancer of  
 CC differing aggressiveness and metastatic potential

XX SQ Sequence 2326 BP; 475 A; 690 C; 625 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 3; Length 2326;

Best Local Similarity 100.0%; Pred. No. 0.71;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCATGTTACAGGTAGAAAAGCC 22

Db 1847 GCATGTTACAGGTAGAAAAGCC 1826

RESULT 12

ADB75506/c

ID ADB75506 standard; cDNA; 2326 BP.

XX AC ADB75506;

XX DT 04-DEC-2003 (first entry)

XX DE Prostate cancer marker cDNA.

XX KW Prostate; cancer; cytostatic; gene therapy; marker; ss.

XX OS Homo sapiens.

XX FN WO2003009814-A2.

XX PD 06-FEB-2003.

XX PF 25-JUL-2002; 2002WO-US023913.

XX PR 25-JUL-2001; 2001US-0307982P.

XX PR 22-AUG-2001; 2001US-0314356P.

XX PR 25-SEP-2001; 2001US-0325020P.

XX PR 12-DEC-2001; 2001US-0341746P.

XX PR 05-MAR-2002; 2002US-0362158P.

XX PA (MTLL-) MILLENNIUM PHARM INC.

XX PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

XX Hoersh S, Kamatkar S, Wonsay AM, Glatt K, Zhao X, Anderson D;

XX WPI; 2003-248033/24.

XX DR New nucleic acid molecule, useful for diagnosing or treating prostate  
 XX cancer.

XX PS Disclosure; SEQ ID NO 330; 99pp; English.

XX CC The invention relates to newly discovered cancer markers associated with  
 CC the cancerous state of prostate cells. Also disclosed is a method of  
 CC assessing whether a patient is afflicted with prostate cancer. The method  
 CC of the invention involves assessing whether a patient is afflicted with  
 CC prostate cancer by comparing the level of expression of a marker in a  
 CC patient sample and the normal level of expression of the marker in a  
 CC control non-prostate cancer sample, where a significant increase in the  
 CC level of expression of the marker in the patient sample and the normal  
 CC level indicates that the patient is afflicted with prostate cancer.  
 CC Nucleic acids of the invention are useful for diagnosing or treating  
 CC prostate cancer, and may be useful in gene therapy. Sequences given in  
 CC ADB75177-ADB75611 represent marker cDNA and proteins. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at



CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 2326 BP; 475 A; 690 C; 625 G; 536 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 22; DB 9; Length 2326;

XX Best Local Similarity 100.0%; Pred. No. 0.71;

SQ Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22

|||||  
1847 GCATGTTACAGGTAGAAAAGCC 1826

DB

RESULT 13

AAV49590/c

ID AAV49590 standard; cDNA to mRNA; 2373 BP.

XX AC AAV49590;

XX DT 21-OCT-1998 (first entry)

XX DE Human liver cell clone HP10302 cDNA #1.

XX KW Transmembrane domain; human; nutrition; cytokine; cell proliferation;

XX KW differentiation; immune system; stimulator; suppressor; regulator;

XX KW hematopoiesis; activin; inhibitor; chemotactic; chemokinetic; receptor;

XX KW haemostatic; thrombolytic; ligand; anti-inflammatory; tumour; ds.

XX OS Homo sapiens.

XX FH Key

FT CDS Location/Qualifiers

FT 134..1813

FT /\*tag= a

FT /product= "transmembrane domain containing protein"

XX WO9821328-A2.

XX PD 22-MAY-1998.

XX PF 07-NOV-1997; 97WO-JP004056.

XX PR 13-NOV-1996; 96JP-00301429.

XX PA (SAGA) SAGAMI CHEM RES CENTRE.

XX PA (PROT-) PROTEGENE INC.

XX PI Kato S, Sekine S, Yamaguchi T, Kobayashi M;

XX DR WPI; 1998-297932/26.

XX DR P-PSDB; AAW64554.

XX PT Human protein having transmembrane domain - useful for, e.g. research and

XX PT nutrition.

XX PS Claim 4; Page 166-169; 205pp; English.

XX CC AAV49550-V49599 are cDNA sequences which encode human proteins containing

XX CC a transmembrane domain. These proteins can be used for, e.g. research and

XX CC nutrition, and may have cytokine and cell proliferation/differentiation,

XX CC immune stimulating/suppressing, hematopoiesis regulating, tissue growth,

XX CC activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic,

XX CC receptor/ligand, anti-inflammatory or tumour inhibition activity

XX SQ Sequence 2373 BP; 469 A; 716 C; 645 G; 543 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 22; DB 2; Length 2373;

XX Best Local Similarity 100.0%; Pred. No. 0.71;

XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22

|||||  
1904 GCATGTTACAGGTAGAAAAGCC 1883

DB

RESULT 14

AAL05398

ID AAL05398 standard; DNA; 3160 BP.

XX AC AAL05398;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen DNA SEQ ID NO: 8086.

XX KW Human; reproductive system related antigen; reproductive system disorder;

XX KW cancer; gene therapy; ds.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001339.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 14-AUG-2000; 2000US-0225759P.

XX PR 18-AUG-2000; 2000US-0226681P.

XX PR 22-AUG-2000; 2000US-0226681P.

XX PR 22-AUG-2000; 2000US-0226686P.

XX PR 22-AUG-2000; 2000US-0227182P.

XX PR 23-AUG-2000; 2000US-0227009P.

XX PR 30-AUG-2000; 2000US-0228924P.

XX PR 01-SEP-2000; 2000US-0228928P.

XX PR 01-SEP-2000; 2000US-0229343P.

XX PR 01-SEP-2000; 2000US-0229344P.

XX PR 01-SEP-2000; 2000US-0229345P.

XX PR 05-SEP-2000; 2000US-0229509P.

XX PR 05-SEP-2000; 2000US-0229513P.

XX PR 06-SEP-2000; 2000US-0230437P.

XX PR 06-SEP-2000; 2000US-0230438P.

XX PR 08-SEP-2000; 2000US-0231242P.

XX PR 08-SEP-2000; 2000US-0231243P.

XX PR 08-SEP-2000; 2000US-0231244P.

XX PR 08-SEP-2000; 2000US-0231413P.

XX PR 08-SEP-2000; 2000US-0231414P.

XX PR 08-SEP-2000; 2000US-0232080P.

XX PR 08-SEP-2000; 2000US-0232081P.





XX  
SQ Sequence 3160 BP; 861 A; 761 C; 855 G; 683 T; 0 U; 0 Other;  
Query Match 100.0%; Score 22; DB 4; Length 3160;  
Best Local Similarity 100.0%; Pred. No. 0.74;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCATGTTACAGGTAGAAAGCC 22  
|||  
Db 945 GCATGTTACAGGTAGAAAGCC 966

Search completed: April 10, 2004, 20:05:47  
Job time : 11.2273 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 14:52:03 ; Search time 2.03516 Seconds  
(without alignments)  
5998.994 Million cell updates/sec

Title: US-09-743-825-7  
Perfect score: 22  
Sequence: 1 gcatgtacaggtagaaagcc 22

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/PCFUS\_COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfileseq1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17.4	79.1	478	US-09-621-976-1977	Sequence 1977, Ap
C 2	17.2	78.2	1559	US-09-602-472A-3	Sequence 3, Appli
C 3	16.8	76.4	1098	US-09-134-001C-2238	Sequence 2238, Ap
C 4	16.2	73.6	6568	US-08-957-636-1	Sequence 1, Appli
C 5	16	72.7	4615	US-08-674-351-3	Sequence 3, Appli
C 6	15.8	71.8	747	US-08-750-145A-19	Sequence 19, Appl
C 7	15.8	71.8	747	US-08-975-698A-23	Sequence 23, Appl
C 8	15.8	71.8	747	US-09-417-090-23	Sequence 23, Appl
C 9	15.8	71.8	747	US-09-727-578-23	Sequence 23, Appl
C 10	15.8	71.8	1671	US-09-023-655-1034	Sequence 1034, Ap
C 11	15.8	71.8	7904	US-08-316-239B-1	Sequence 1, Appli
C 12	15.8	71.8	7904	US-08-316-239B-2	Sequence 2, Appli
C 13	15.8	71.8	7904	US-08-410-005-1	Sequence 1, Appli
C 14	15.8	71.8	7904	US-08-929-140-1	Sequence 1, Appli
C 15	15.8	71.8	7904	US-09-560-579A-1	Sequence 1, Appli
C 16	15.8	71.8	14051	US-08-956-171B-103	Sequence 103, App
C 17	15.8	71.8	53332	US-09-801-861-3	Sequence 3, Appli
C 18	15.6	70.9	523	US-09-833-381-281	Sequence 281, App
C 19	15.6	70.9	1593	US-09-602-472A-1	Sequence 1, Appli
C 20	15.6	70.9	1740	US-08-471-025-37	Sequence 37, Appl
C 21	15.6	70.9	2196	US-09-149-476-289	Sequence 289, App
C 22	15.6	70.9	2842	US-09-620-312D-574	Sequence 574, App
C 23	15.6	70.9	3454	US-09-620-312D-167	Sequence 167, App
C 24	15.6	70.9	5402	US-09-620-312D-194	Sequence 194, App
C 25	15.6	70.9	9391	US-09-562-702A-11	Sequence 11, Appl
C 26	15.6	70.9	9511	US-09-562-702A-9	Sequence 9, Appli
C 27	15.6	70.9	99916	US-09-816-095-3	Sequence 3, Appli

C 28	15.4	70.0	426	4	US-09-621-976-11345	Sequence 11345, A
C 29	15.4	70.0	2475	2	US-07-684-965-5	Sequence 5, Appli
C 30	15.4	70.0	2475	3	US-09-252-658-5	Sequence 5, Appli
C 31	15.4	70.0	161652	4	US-09-497-855A-40	Sequence 40, Appl
C 32	15.2	69.1	20	4	US-09-198-452A-5333	Sequence 5333, Ap
C 33	15.2	69.1	315	3	US-09-157-177-104	Sequence 104, App
C 34	15.2	69.1	1014	4	US-09-328-352-3704	Sequence 3704, App
C 35	15.2	69.1	4078	4	US-09-016-434-870	Sequence 870, App
C 36	15.2	69.1	4371	4	US-09-134-000C-3022	Sequence 3022, Ap
C 37	15.2	69.1	4331	3	US-08-726-320-2	Sequence 2, Appli
C 38	15.2	69.1	4331	3	US-09-208-716-2	Sequence 2, Appli
C 39	15.2	69.1	5924	4	US-08-956-171B-130	Sequence 130, App
C 40	15.2	69.1	6525	1	US-08-493-092-3	Sequence 3, Appli
C 41	15.2	69.1	6525	1	US-08-508-836A-3	Sequence 3, Appli
C 42	15.2	69.1	7158	4	US-09-543-681A-2132	Sequence 2132, Ap
C 43	15.2	69.1	11970	3	US-09-345-217-1	Sequence 1, Appli
C 44	15.2	69.1	13205	4	US-09-835-811-3	Sequence 3, Appli
C 45	15.2	69.1	15567	4	US-09-627-376-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-621-976-1977/c  
; Sequence 1977, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 1977  
; LENGTH: 478  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 203..433  
; NAME/KEY: sig\_peptide  
; LOCATION: 203..397  
; OTHER INFORMATION: Von Heijne matrix  
; OTHER INFORMATION: score 3.70000004768372  
; OTHER INFORMATION: seq LTFSLGLSPIYV/RR  
US-09-621-976-1977

Query Match 79.1%; Score 17.4; DB 4; Length 478;

Best Local Similarity 94.7%; Pred. No. 18;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 TGTTCAGGTAGAAAAGCC 22  
DB 421 TGTTCAGGTAGAAAAGCC 403

RESULT 2

US-09-602-472A-3  
; Sequence 3, Application US/09602472A  
; Patent No. 6608240  
; GENERAL INFORMATION:  
; APPLICANT: Bidney, Dennis L.  
; APPLICANT: Hu, Xu  
; APPLICANT: Li, Guihua  
; TITLE OF INVENTION: Sunflower Disease Resistance Genes  
; FILE REFERENCE: 35718/200630  
; CURRENT APPLICATION NUMBER: US/09/602,472A  
; CURRENT FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/140,876

```
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1559
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)...(1319)
US-09-602-472A-3

Query Match          78.2%; Score 17.2; DB 4; Length 1559;
Best Local Similarity 86.4%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCATGTTACAGGTAGAAAAGCC 22
Db 917 GCTTATTACCGGTAGAAAAGCC 938

RESULT 3
US-09-134-001C-2238
; Sequence 2238, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2238
; LENGTH: 1098
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2238

Query Match          76.4%; Score 16.8; DB 4; Length 1098;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCATGTTACAGGTAGAAAAG 20
Db 891 GCATAGTACAGGTAGAAAAG 910

RESULT 4
US-08-857-636-1
; Sequence 1, Application US/08857636
; Patent No. 6552181
; GENERAL INFORMATION:
; APPLICANT: Dean, Michael Carlton
; APPLICANT: Hann, Heidi Eve
; APPLICANT: Wickling, Carol
; APPLICANT: Christiansen, Jeffrey
; APPLICANT: Zaphiropoulos, Peter G.
; APPLICANT: Galani, Mae R.
; APPLICANT: Shanley, Susan Mary
; APPLICANT: Chidambaram, Abirami
; APPLICANT: Vorechovsky, Igor
; APPLICANT: Holmberg-Lindstrom, Erika
; APPLICANT: Unden, Anne Birgitte
; APPLICANT: Gillies, Susan Alana
; APPLICANT: Negus, Kylie
; APPLICANT: Smyth, Ian Mcleod
; APPLICANT: Pressman, Carol Leah
; APPLICANT: Leffell, David J.
```

```
; APPLICANT: Gerrard, Bernard
; APPLICANT: Goldstein, Alisa Miriam
; APPLICANT: Mainwright, Brandon
; APPLICANT: Toftgard, Rune Carl-Magnus
; APPLICANT: Chenevix-Trench, Georgia
; APPLICANT: Bale, Allen E.
; TITLE OF INVENTION: A Basal Cell Carcinoma Tumor Suppressor Gene
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/857,636
; FILING DATE: 16-MAY-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,906
; FILING DATE: 17-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P00011
; FILING DATE: 21-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P00363
; FILING DATE: 07-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/019,765
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 015280-27820005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..6568
; OTHER INFORMATION: /note="human nevoid basal cell
; OTHER INFORMATION: carcinoma syndrome (NBCCS)
; OTHER INFORMATION: (PATCHED (PTC)) cDNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 442..4332
US-08-857-636-1

Query Match          73.6%; Score 16.2; DB 4; Length 6568;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CATGTTACAGGTAGAAAAGCC 22
Db 5454 CCTGTCACAGGTAGAAAACC 5474

RESULT 5
US-08-674-351-3
; Sequence 3, Application US/08674351
```

Patent No. 5831013  
GENERAL INFORMATION:  
APPLICANT: Bruenn, Jeremy A.  
APPLICANT: Yao, Wenheng  
TITLE OF INVENTION: CAPSID POLYPEPTIDES AND USE TO INHIBIT  
TITLE OF INVENTION: VIRAL PACKAGING  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/674,351  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Timian, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19226/740  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 263-1636  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4615 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
US-08-674-351-3

Query Match 72.7%; Score 16; DB 2; Length 4615;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GTTACAGGTAGAAAAG 20  
DB 1626 GTTACAGGTAGAAAAG 1641

RESULT 6  
US-08-750-145A-19/c  
Sequence 19, Application US/08750145A  
Patent No. 6010851  
GENERAL INFORMATION:  
APPLICANT: MIHARA, Yasuhiro  
APPLICANT: MIHARA, Takashi  
APPLICANT: YAMADA, Hideaki  
APPLICANT: ASANO, Yasuhisa  
TITLE OF INVENTION: Method for Producing Nucleoside-5'-  
TITLE OF INVENTION: Phosphate Ester  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.  
STREET: 1755 S. JEFF. DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/750,145A  
FILING DATE: 01-JAN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-149781  
FILING DATE: 05-May-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-094680  
FILING DATE: 26-Mar-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: NORMAN F. OBLON  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 0010-0830-0 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 747 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterobacter aerogenes  
STRAIN: IFO 12010  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..744  
US-08-750-145A-19

Query Match 71.8%; Score 15.8; DB 3; Length 747;  
Best Local Similarity 89.5%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGTTCAGGTAGAAAAGCC 22  
DB 454 TGTTCAGGTAGAAAAGCC 436

RESULT 7  
US-08-975-698A-23/c  
Sequence 23, Application US/08975698A  
Patent No. 6015697  
GENERAL INFORMATION:  
APPLICANT: MIHARA, YASUHIRO  
APPLICANT: UTAGAWA, TAKASHI  
APPLICANT: YAMADA, HIDEAKI  
APPLICANT: ASANO, YASUHISA  
TITLE OF INVENTION: METHOD FOR PRODUCING NUCLEOSIDE-5'-PHOSPHATE  
TITLE OF INVENTION: ESTER  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/975,698A  
FILING DATE: 21-NOV-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F  
REGISTRATION NUMBER: 24,618





NAME/KEY: CDS  
LOCATION: 1..744  
US-09-727-578-23

Query Match 71.8%; Score 15.8; DB 4; Length 747;  
Best Local Similarity 89.5%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGTTACAGGTAGAAAGCC 22  
|||||  
DB 454 TGTTACAGGTAGAAAGCC 436

RESULT 10  
US-09-023-655-1034  
Sequence 1034, Application US/09023655  
Patent No. 6607879  
GENERAL INFORMATION:  
APPLICANT: Cocks, Benjamin G.  
APPLICANT: Susan G. Stuart  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,655  
FILING DATE: HEREMITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1034:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1671 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g180152

US-09-023-655-1034  
Query Match 71.8%; Score 15.8; DB 4; Length 1671;  
Best Local Similarity 89.5%; Pred. No. 1.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ATGTTACAGGTAGAAAGC 21  
|||||  
DB 1220 ATGTTACTGTAGAAAGCC 1238

RESULT 11  
US-08-316-239B-1

Sequence 1, Application US/08316239B  
Patent No. 5679509  
GENERAL INFORMATION:  
APPLICANT: Wheeler, Cosette M.  
APPLICANT: Parmenter, Cheryl A.  
TITLE OF INVENTION: Methods and a Diagnostic Aid for  
TITLE OF INVENTION: Distinguishing a Subset of HPV that is Associated with an  
TITLE OF INVENTION: Increased Risk of Developing Cervical Dysplasia and  
TITLE OF INVENTION: Cervical Cancer  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jagtiani & Associates  
STREET: 6126 Rocky Way Court  
CITY: Centreville  
STATE: VA  
COUNTRY: USA  
ZIP: 20120-3400  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/316,239B  
FILING DATE: 30-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jagtiani, Ajay A.  
REGISTRATION NUMBER: 35,205  
REFERENCE/DOCKET NUMBER: UNME-0001  
TELEPHONE: (703) 817-9453  
TELEFAX: (703) 803-9387  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7904 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
US-08-316-239B-1

Query Match 71.8%; Score 15.8; DB 1; Length 7904;  
Best Local Similarity 89.5%; Pred. No. 1.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ATGTTACAGGTAGAAAGC 21  
|||||  
DB 1293 ATGTTACAGGTAGAAAGGC 1311

RESULT 12  
US-08-316-239B-2  
Sequence 2, Application US/08316239B  
Patent No. 5679509  
GENERAL INFORMATION:  
APPLICANT: Wheeler, Cosette M.  
APPLICANT: Parmenter, Cheryl A.  
TITLE OF INVENTION: Methods and a Diagnostic Aid for  
TITLE OF INVENTION: Distinguishing a Subset of HPV that is Associated with an  
TITLE OF INVENTION: Increased Risk of Developing Cervical Dysplasia and  
TITLE OF INVENTION: Cervical Cancer  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jagtiani & Associates  
STREET: 6126 Rocky Way Court  
CITY: Centreville  
STATE: VA  
COUNTRY: USA  
ZIP: 20120-3400  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jagtiani, Ajay A.
; REGISTRATION NUMBER: 35,205
; REFERENCE/DOCKET NUMBER: UNME-0001
; TELEPHONE: (703) 817-9453
; TELEFAX: (703) 803-9387
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7904 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-316-239B-2

```

```

Query Match 71.8%; Score 15.8; DB 1; Length 7904;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ATGTTACAGGTAGAAAGC 21
Db 1293 ATGTTACAGGTAGAAAGGC 1311

```

```

RESULT 13
US-08-410-005-1
; Sequence 1, Application US/08410005
; Patent No. 5683902
; GENERAL INFORMATION:
; APPLICANT: Hampel, Arnold
; APPLICANT: DiPaolo, Joseph
; APPLICANT: Siwkowski, Andrew M.
; APPLICANT: Galasinski, Scott C.
; TITLE OF INVENTION: HUMAN PAPILLOMA VIRUS INHIBITION BY A
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard, Perry & Milton
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: MI
; COUNTRY: USA
; ZIP: 48099-9998
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 21-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 08/242,665
; FILING DATE: 13-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-301 (NIU)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 7904 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Human papillomavirus
; STRAIN: HPV16
US-08-410-005-1

Query Match 71.8%; Score 15.8; DB 1; Length 7904;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ATGTTACAGGTAGAAAGC 21
Db 1293 ATGTTACAGGTAGAAAGGC 1311

RESULT 14
US-08-929-140-1
; Sequence 1, Application US/08929140
; Patent No. 6084090
; GENERAL INFORMATION:
; APPLICANT: DiPaolo, Joseph
; APPLICANT: Alvarez-Salas, Luis
; TITLE OF INVENTION: HUMAN PAPILLOMA VIRUS INHIBITION
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive Sixteenth Flo
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,140
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: NIH138.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714/760-0404
; TELEFAX: 714/760-9503
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7904 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-929-140-1

Query Match 71.8%; Score 15.8; DB 3; Length 7904;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ATGTTACAGGTAGAAAGC 21

```

Db 1293 ATGTTACAGGTAGAGGC 1311

RESULT 15

US-09-560-579A-1  
 ; Sequence 1, Application US/09560579A  
 ; Patent No. 6277980  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DiPaolo, Joseph  
 ; ; Alvarez-Salas, Luis  
 ; TITLE OF INVENTION: HUMAN PAPILLOMA VIRUS INHIBITION  
 ; ; BY ANTISENSE OLIGONUCLEOTIDES  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Knobbe, Martens, Olson & Bear  
 ; STREET: 620 Newport Center Drive Sixteenth Flo  
 ; CITY: Newport Beach  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 92660  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Dikette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/560,579A  
 ; FILING DATE: 28-Apr-2000  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/929,140  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Altman, Daniel E  
 ; REGISTRATION NUMBER: 34,115  
 ; REFERENCE/DOCKET NUMBER: NIH138.001A  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 714/760-0404  
 ; TELEFAX: 714/760-9503  
 ; TELEX: <Unknown>  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 7904 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: Genomic DNA  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 ; US-09-560-579A-1

Query Match 71.8% Score 15.8; DB 3; Length 7904;  
 Best Local Similarity 89.5%; Pred.No. 1.7e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ATGTTACAGGTAGAGGC 21  
 |||||  
 Db 1293 ATGTTACAGGTAGAGGC 1311

Search completed: April 10, 2004, 20:25:39  
 Job time : 3.03516 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 14:29:57 ; Search time 8.66555 Seconds  
(without alignments)  
9524.167 Million cell updates/sec

Title: US-09-743-825-7

Perfect score: 22

Sequence: 1 gcattacaggtagaagcc 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 247585 seqs, 1875730760 residues

Total number of hits satisfying chosen parameters: 4951170

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*

14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*

17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*

18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	100.0	366	14	US-10-066-543-1782
C 2	22	100.0	457	9	US-09-880-107-8
C 3	22	100.0	856	10	US-09-764-891-1557
C 4	22	100.0	1737	10	US-09-764-891-8087
C 5	22	100.0	2105	12	US-10-262-839-127
C 6	22	100.0	2233	15	US-10-074-978A-83
C 7	22	100.0	2284	12	US-10-262-839-129
C 8	22	100.0	2326	12	US-10-342-887-932
C 9	22	100.0	2326	14	US-10-205-823-330
C 10	22	100.0	2373	10	US-09-284-320-71
C 11	22	100.0	3160	10	US-09-764-891-8086
C 12	22	100.0	7709	10	US-09-764-891-8085
C 13	22	100.0	143899	10	US-09-972-546-15
C 14	17.4	79.1	266	14	US-10-007-280A-4
C 15	17.4	79.1	1483	14	US-10-007-280A-5

C 16	17.4	79.1	1630	12	US-09-823-245A-623	Sequence 623, App
C 17	17.4	79.1	2005	15	US-10-104-047-288	Sequence 288, App
C 18	17.2	78.2	2939	15	US-10-320-797-58	Sequence 58, Appl
C 19	17	77.3	710	15	US-10-027-632-21770	Sequence 21770, A
C 20	16.8	76.4	637	15	US-10-027-632-247072	Sequence 247072, A
C 21	16.8	76.4	810	15	US-10-027-632-7665	Sequence 7665, Ap
C 22	16.8	76.4	1089	12	US-10-282-122A-34990	Sequence 34990, A
C 23	16.8	76.4	1472	12	US-10-424-599-95670	Sequence 95670, A
C 24	16.8	76.4	1719	12	US-10-424-599-6371	Sequence 6371, Ap
C 25	16.4	74.5	448	10	US-09-814-353-14497	Sequence 14497, A
C 26	16.4	74.5	502	10	US-09-814-353-1763	Sequence 1763, Ap
C 27	16.4	74.5	502	10	US-09-814-353-8113	Sequence 8113, Ap
C 28	16.4	74.5	565	15	US-10-027-632-228420	Sequence 228420, A
C 29	16.4	74.5	577	15	US-10-027-632-278135	Sequence 278135, A
C 30	16.4	74.5	619	15	US-10-027-632-278136	Sequence 278136, A
C 31	16.4	74.5	935	12	US-10-425-114-28559	Sequence 28559, A
C 32	16.4	74.5	1522	15	US-10-027-632-264145	Sequence 264145, A
C 33	16.4	74.5	2355	12	US-10-282-122A-41960	Sequence 41960, A
C 34	16.4	74.5	2397	15	US-10-369-493-44910	Sequence 44910, A
C 35	16.4	74.5	3587	16	US-10-389-566-193	Sequence 193, App
C 36	16.4	74.5	4911	10	US-09-814-353-19968	Sequence 19968, A
C 37	16.2	73.6	265	12	US-10-085-783A-42740	Sequence 42740, A
C 38	16.2	73.6	265	15	US-10-242-535A-42740	Sequence 42740, A
C 39	16.2	73.6	490	12	US-10-085-783A-30540	Sequence 30540, A
C 40	16.2	73.6	490	15	US-10-242-535A-30540	Sequence 30540, A
C 41	16.2	73.6	495	15	US-10-027-632-61280	Sequence 61280, A
C 42	16.2	73.6	532	15	US-10-027-632-280485	Sequence 280485, A
C 43	16.2	73.6	551	11	US-09-864-408A-2691	Sequence 2691, Ap
C 44	16.2	73.6	556	14	US-10-023-386-6433	Sequence 6433, Ap
C 45	16.2	73.6	568	15	US-10-027-632-59334	Sequence 59334, A

ALIGNMENTS

RESULT 1

US-10-066-543-1782/c  
; Sequence 1782, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuhui  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1782  
; LENGTH: 366  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-066-543-1782

Query Match 100.0%; Score 22; DB 14; Length 366;  
Best Local Similarity 100.0%; Pred. No. 0.72;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATTTACAGGTAGAAAGCC 22

Db 150 GCATTTACAGGTAGAAAGCC 129

```
RESULT 2
US-09-880-107-8
; Sequence 8, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA004521
US-09-880-107-8

Query Match      100.0%; Score 22; DB 9; Length 457;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
Db      44 GCATGTTACAGGTAGAAAAGCC 65

RESULT 3
US-09-764-891-1557
; Sequence 1557, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1557
; LENGTH: 856
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-1557

Query Match      100.0%; Score 22; DB 10; Length 856;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
Db      63 GCATGTTACAGGTAGAAAAGCC 84

RESULT 4
US-09-764-891-8087
; Sequence 8087, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8087
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8087

Query Match      100.0%; Score 22; DB 10; Length 1737;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      . 1 GCATGTTACAGGTAGAAAAGCC 22
Db      944 GCATGTTACAGGTAGAAAAGCC 965

RESULT 5
US-10-262-839-127/c
; Sequence 127, Application US/10262839
; Publication No. US20040038877A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John,
; APPLICANT: Anderson, David W.,
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Catteston, Elina,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Leach, Martin,
; APPLICANT: Li, Li,
; APPLICANT: Miller, Charles,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Vernier, Corine,
; APPLICANT: Voss, Edward,
; APPLICANT: Zerhusen, Brian,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462A
; CURRENT APPLICATION NUMBER: US/10/262,839
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
```

```
/ PRIOR FILING DATE: 2001-10-09
/ PRIOR APPLICATION NUMBER: 60/328,849
/ PRIOR FILING DATE: 2001-10-12
/ PRIOR APPLICATION NUMBER: 60/374,738
/ PRIOR FILING DATE: 2002-04-23
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 367
/ SOFTWARE: CuraseqList version 0.1
/ SEQ ID NO 127
/ LENGTH: 2105
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (47)..(1588)
US-10-262-839-127

Query Match      100.0%; Score 22; DB 12; Length 2105;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
      |||||
Db      1682 GCATGTTACAGGTAGAAAAGCC 1661

RESULT 6
US-10-074-978A-83/c
/ Sequence 83, Application US/10074978A
/ Publication No. US20040010119A1
/ GENERAL INFORMATION:
/ APPLICANT: Leite, Mario
/ APPLICANT: Spytek, Kimberly A
/ APPLICANT: Guo, Xiaojia (Sasha)
/ APPLICANT: Fernandes, Elma
/ APPLICANT: Li, Li
/ APPLICANT: Kekuda, Ramesh
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Casman, Stacie
/ APPLICANT: Boldog, Ferenc
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Blalock, Angela
/ APPLICANT: Ballinger, Robert
/ APPLICANT: Vernet, Corine
/ APPLICANT: Tchernev, Velizar T
/ APPLICANT: Malyankar, Uriel M
/ APPLICANT: Gusev, Vladimir
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Mezes, Peter S
/ APPLICANT: Ellerman, Karen
/ APPLICANT: Heyes, Melvin P
/ APPLICANT: Herrman, John
/ APPLICANT: Pena, Carol E A
/ APPLICANT: Shimkete, Richard A
/ APPLICANT: Taupier Jr, Raymond J
/ APPLICANT: Moore, No. US20040010119A1lle
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Edinger, Shlomit
/ APPLICANT: Gunther, Erik
/ APPLICANT: Stone, Dave
/ APPLICANT: Millet, Isabelle
/ APPLICANT: Peyman, John
/ APPLICANT: Smithson, Glenna
/ TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 21402-269
/ CURRENT APPLICATION NUMBER: US/10/074,978A
/ CURRENT FILING DATE: 2003-01-07
/ PRIOR APPLICATION NUMBER: 60/368,221
/ PRIOR FILING DATE: 2001-02-12
/ PRIOR APPLICATION NUMBER: 60/335,109
/ PRIOR FILING DATE: 2001-10-31
/ PRIOR APPLICATION NUMBER: 60/312,284
/ PRIOR FILING DATE: 2001-08-14
```

```
/ PRIOR APPLICATION NUMBER: 60/268,496
/ PRIOR FILING DATE: 2001-02-13
/ PRIOR APPLICATION NUMBER: 60/276,703
/ PRIOR FILING DATE: 2001-03-16
/ PRIOR APPLICATION NUMBER: 60/330,293
/ PRIOR FILING DATE: 2001-10-18
/ PRIOR APPLICATION NUMBER: 60/322,127
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: 60/280,899
/ PRIOR FILING DATE: 2001-04-02
/ PRIOR APPLICATION NUMBER: 60/310,797
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: 60/268,646
/ PRIOR FILING DATE: 2001-02-14
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 547
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 83
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-074-978A-83

Query Match      100.0%; Score 22; DB 15; Length 2233;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
      |||||
Db      1754 GCATGTTACAGGTAGAAAAGCC 1733

RESULT 7
US-10-262-839-129/c
/ Sequence 129, Application US/10262839
/ Publication No. US20040038877A1
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook, John,
/ APPLICANT: Anderson, David W.,
/ APPLICANT: Boldog, Ferenc,
/ APPLICANT: Burgess, Catherine,
/ APPLICANT: Catterton, Elina,
/ APPLICANT: Edinger, Shlomit,
/ APPLICANT: Ellerman, Karen,
/ APPLICANT: Gerlach, Valerie,
/ APPLICANT: Gorman, Linda,
/ APPLICANT: Guo, Xiaojia,
/ APPLICANT: Ji, Weizhen,
/ APPLICANT: Kekuda, Ramesh,
/ APPLICANT: Leach, Martin,
/ APPLICANT: Li, Li,
/ APPLICANT: Miller, Charles,
/ APPLICANT: Patturajan, Meera,
/ APPLICANT: Reiger, Daniel,
/ APPLICANT: Rothenberg, Mark,
/ APPLICANT: Shimkets, Richard,
/ APPLICANT: Smithson, Glenna,
/ APPLICANT: Spytek, Kimberly, Jr.,
/ APPLICANT: Taupier, Raymond, Jr.,
/ APPLICANT: Vernet, Corine,
/ APPLICANT: Voss, Edward,
/ APPLICANT: Zerhusen, Brian,
/ APPLICANT: Zhong, Mei
/ TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
/ FILE REFERENCE: 21402-462A
/ CURRENT APPLICATION NUMBER: US/10/262,839
/ CURRENT FILING DATE: 2002-10-01
/ PRIOR APPLICATION NUMBER: 60/326,483
/ PRIOR FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: 60/327,917
/ PRIOR FILING DATE: 2001-10-09
/ PRIOR APPLICATION NUMBER: 60/328,029
/ PRIOR FILING DATE: 2001-10-09
```

```
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 129
; LENGTH: 2284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1768)
US-10-262-839-129

Query Match      100.0%; Score 22; DB 12; Length 2284;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
      |||
Db      1862 GCATGTTACAGGTAGAAAAGCC 1841

RESULT 8
US-10-342-887-932/c
; Sequence 932, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 932
; LENGTH: 2326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-932

Query Match      100.0%; Score 22; DB 12; Length 2326;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
      |||
Db      1847 GCATGTTACAGGTAGAAAAGCC 1826
```

```
RESULT 9
US-10-205-823-330/c
; Sequence 330, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsay, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 2326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-330

Query Match      100.0%; Score 22; DB 14; Length 2326;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
      |||
Db      1847 GCATGTTACAGGTAGAAAAGCC 1826

RESULT 10
US-09-284-320-71/c
; Sequence 71, Application US/09284320
; Publication No. US20030092175A1
; GENERAL INFORMATION:
; APPLICANT: Kato, Seishi et al.
; TITLE OF INVENTION: HUMAN PROTEINS HAVING TRANSMEMBRANE DOMAINS AND DNAs
; TITLE OF INVENTION: ENCODING THESE PROTEINS
; FILE REFERENCE: GIN-6705CPUS
; CURRENT APPLICATION NUMBER: US/09/284,320
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: JP 8-301429
; PRIOR FILING DATE: 1996-11-13
; PRIOR APPLICATION NUMBER: PCT/JP97/04056
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 2373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```



```

; LOCATION: (134)..(1810)
US-09-284-320-71

Query Match      100.0%; Score 22; DB 10; Length 2373;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
      |||
Db      1904 GCATGTTACAGGTAGAAAAGCC 1883

RESULT 11
US-09-764-891-8086
; Sequence 8086, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8086
; LENGTH: 3160
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8086

Query Match      100.0%; Score 22; DB 10; Length 3160;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
      |||
Db      945 GCATGTTACAGGTAGAAAAGCC 966

RESULT 12
US-09-764-891-8085
; Sequence 8085, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8085
; LENGTH: 7709
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8085

Query Match      100.0%; Score 22; DB 10; Length 7709;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
      |||
Db      943 GCATGTTACAGGTAGAAAAGCC 964

RESULT 13
US-09-972-546-15
; Sequence 15, Application US/09972546
; Publication No. US20030124704A1
; GENERAL INFORMATION:

```

```

; APPLICANT: STRITTMATTER, STEPHEN M.
; APPLICANT: CAE, RICHARD L.
; APPLICANT: SAH, DINAH W. Y.
; TITLE OF INVENTION: NOCO RECEPTOR HOMOLOGS
; FILE REFERENCE: ALI6US
; CURRENT APPLICATION NUMBER: US/09/972,546
; CURRENT FILING DATE: 2001-10-06
; PRIOR APPLICATION NUMBER: 60/238,361
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 143899
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2044)..(2144)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (6609)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (6625)..(6724)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (14153)..(14252)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (19512)..(19611)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (22595)..(22694)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (27825)..(27924)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (34953)..(35052)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (40783)..(40882)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (49000)..(49099)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (62884)..(62983)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (75528)..(75627)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (87944)..(88043)
; OTHER INFORMATION: a, t, c, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (111030)..(111129)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-09-972-546-15

Query Match      100.0%; Score 22; DB 10; Length 143899;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCATGTTACAGGTAGAAAAGCC 22
      |||
Db      38486 GCATGTTACAGGTAGAAAAGCC 38507

RESULT 14
US-10-007-280A-4
; Sequence 4, Application US/10007280A
; Publication No. US20030059784A1

```

GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Herve  
; APPLICANT: Salceda, Susana  
; APPLICANT: Chenghua, Liu  
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro  
; FILE REFERENCE: DEX-0257  
; CURRENT APPLICATION NUMBER: US/10/007,280A  
; CURRENT FILING DATE: 2001-11-07  
; PRIOR APPLICATION NUMBER: US 60/246,640  
; PRIOR FILING DATE: 2000-11-08  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 266  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-007-280A-4

Query Match 79.1%; Score 17.4; DB 14; Length 266;  
Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ATGTTACAGGTAGAAAAGC 21  
|||||  
Db 168 ATGTTACAGGTACAAAGC 186

RESULT 15  
US-10-007-280A-5  
; Sequence 5, Application US/10007280A  
; Publication No. US20030059784A1  
; GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Herve  
; APPLICANT: Salceda, Susana  
; APPLICANT: Chenghua, Liu  
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro  
; FILE REFERENCE: DEX-0257  
; CURRENT APPLICATION NUMBER: US/10/007,280A  
; CURRENT FILING DATE: 2001-11-07  
; PRIOR APPLICATION NUMBER: US 60/246,640  
; PRIOR FILING DATE: 2000-11-08  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1483  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-007-280A-5

Query Match 79.1%; Score 17.4; DB 14; Length 1483;  
Best Local Similarity 94.7%; Pred. No. 1.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ATGTTACAGGTAGAAAAGC 21  
|||||  
Db 622 ATGTTACAGGTACAAAGC 640

Search completed: April 10, 2004, 20:21:39  
Job time : 9.66555 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:16:47 ; Search time 74.9602 seconds  
(without alignments)  
8764.220 Million cell updates/sec

Title: US-09-743-825-7

Perfect score: 22

Sequence: 1 gcatgtacagtagaagcc 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gas\_hum:\*  
18: em\_gas\_inv:\*  
19: em\_gas\_pln:\*  
20: em\_gas\_vrt:\*  
21: em\_gas\_fun:\*  
22: em\_gas\_mam:\*  
23: em\_gas\_mus:\*  
24: em\_gas\_pro:\*  
25: em\_gas\_rod:\*  
26: em\_gas\_pbg:\*  
27: em\_gas\_vrl:\*  
28: gb\_ges1:\*  
29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	163	9	AA177117 nc03b12.s
2	22	100.0	214	13	BQ301150 CM4-HB002
3	22	100.0	230	13	BQ309353 ILO-BT100
4	22	100.0	230	13	BQ309414 ILO-BT100

5	22	100.0	248	9	AA229970
6	22	100.0	302	10	AW136898
7	22	100.0	344	14	T74926
8	22	100.0	345	9	AA448887
9	22	100.0	420	14	RO0504
10	22	100.0	425	9	AA228376
11	22	100.0	449	10	AW571588
12	22	100.0	457	9	AA004521
13	22	100.0	505	10	AW966858
14	22	100.0	512	9	AA228367
15	22	100.0	522	10	AW247896
16	22	100.0	529	14	CF141328
17	22	100.0	555	13	BQ185119
18	22	100.0	571	13	BX383685
19	22	100.0	582	9	AA176965
20	22	100.0	586	12	BM725383
21	22	100.0	613	12	BM677811
22	22	100.0	617	10	AW956686
23	22	100.0	622	10	AW967482
24	22	100.0	656	12	BQ005970
25	22	100.0	665	12	BQ006067
26	22	100.0	677	13	BX103969
27	22	100.0	726	10	BE872354
28	22	100.0	815	10	BF664173
29	22	100.0	823	10	BG032562
30	22	100.0	834	10	BF673273
31	22	100.0	884	13	BU182024
32	22	100.0	923	10	BE293435
33	22	100.0	1201	9	AL581753
34	21.6	98.2	982	9	AL580746
35	21	95.5	555	12	BQ020153
36	20.4	92.7	837	12	BG497880
37	20	90.9	1171	13	BX447961
38	20	90.9	1201	9	AL583366
39	19.4	88.2	363	28	AZ512665
40	19.4	88.2	536	28	AZ001869
41	19.4	88.2	556	28	AQ997950
42	19.4	88.2	684	13	BU360774
43	19.4	88.2	910	10	BF141223
44	18.8	85.5	685	29	BX158002
45	18.8	85.5	718	14	CB171493

## ALIGNMENTS

RESULT 1  
AA177117  
LOCUS  
DEFINITION  
nc03b12.s1 NCI\_CGAP\_Pr2 Homo sapiens cDNA clone IMAGE:312, mRNA  
163 bp  
linear  
EST 14-AUG-1997  
sequence.  
ACCESSION  
AA177117  
VERSION  
AA177117.1  
KEYWORDS  
GI:1758275  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 163)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)

Seq primer: -40M13 fwd. from Amerham  
High quality sequence stop: 98.

#### FEATURES

Location/Qualifiers  
1. .163  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:312"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr2"  
/note="Vector: pAMP10; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

#### ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 163;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
DB 80 GCATGTTACAGGTAGAAAAGCC 101

#### RESULT 2

BQ301150/c  
LOCUS BQ301150 214 bp mRNA linear EST 16-MAY-2002  
DEFINITION CM4-HB0025-090101-665-d09 HB0025 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BQ301150  
VERSION BQ301150.1 GI:20816672  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 214)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM4t2=CM4-HB0025-090101-665-d09&t3=2001-01-09&t4=1)

Seq primer: puc 18 forward  
High quality sequence stop: 5  
High quality sequence stop: 213.

#### TITLE

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM4t2=CM4-HB0025-090101-665-d09&t3=2001-01-09&t4=1)

Seq primer: puc 18 forward  
High quality sequence stop: 5  
High quality sequence stop: 213.

#### FEATURES

Location/Qualifiers  
1. .214  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HB0025"  
/note="Organ: bocio tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

#### ORIGIN

Query Match 100.0%; Score 22; DB 13; Length 214;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
DB 110 GCATGTTACAGGTAGAAAAGCC 89

#### RESULT 3

BQ309353/c  
LOCUS BQ309353 230 bp mRNA linear EST 16-MAY-2002  
DEFINITION ILO-BT1000-240701-102-f02-827 BT1000 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BQ309353  
VERSION BQ309353.1 GI:20851896  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 230)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=ILO&t2=ILO-BT1000-240701-102-f02-827&t3=2001-07-24&t4=1)

Seq primer: puc 18 forward  
High quality sequence stop: 13  
High quality sequence stop: 229.

#### TITLE

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=ILO&t2=ILO-BT1000-240701-102-f02-827&t3=2001-07-24&t4=1)

#### JOURNAL

MEDLINE  
PUBMED  
COMMENT

#### COMMENT

#### FEATURES

Location/Qualifiers  
1. .230  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="BT1000"  
/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 22; DB 13; Length 230;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
Db 148 GCATGTTACAGGTAGAAAAGCC 127

## RESULT 4

BQ309414/c  
LOCUS BQ309414 230 bp mRNA linear EST 16-MAY-2002  
DEFINITION IL0-BT1000-250701-102-f02 BT1000 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BQ309414  
VERSION BQ309414.1 GI:20851944  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL  
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
PUBMED 20202663  
COMMENT 10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL0&t2=IL0-BT1000-250701-102-f02&t3=2001-07-25&t4=1)

Seq primer: puc 18 forward  
High quality sequence start: 13  
High quality sequence stop: 229.

## FEATURES

source  
1..230  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="BT1000"

/notes="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 100.0%; Score 22; DB 13; Length 230;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
Db 148 GCATGTTACAGGTAGAAAAGCC 127

## RESULT 5

AA229970  
LOCUS AA229970 248 bp mRNA linear EST 21-AUG-1997  
DEFINITION nc37a02.s1 NCI\_CGAP\_Pr2 Homo sapiens cDNA clone IMAGE:1010282, mRNA sequence.

ACCESSION AA229970  
VERSION AA229970.1 GI:1852263  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 248)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D., Michael Emert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/bbrp/image/image.html  
Seq primer: -31m3 fwd. ET from Amersham  
High quality sequence stop: 215.

Location/Qualifiers  
1..248  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1010282"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr2"  
/note="Vector: PAMP10; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into PAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

## FEATURES

source  
1..248  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1010282"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr2"  
/note="Vector: PAMP10; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into PAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr2"  
/note="Vector: PAMP10; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into PAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr2"  
/note="Vector: PAMP10; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into PAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

## ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 248;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
Db 80 GCATGTTACAGGTAGAAAAGCC 101

## RESULT 6

AW136898  
LOCUS AW136898 302 bp mRNA linear EST 29-OCT-1999  
DEFINITION UI-H-B11-adp-b-02-0-UI.s1 NCI\_CGAP\_Sub3 Homo sapiens cDNA clone IMAGE:2717499 3', mRNA sequence.

ACCESSION AW136898  
 VERSION AW136898.1 GI:6141031  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 302)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 TITLE Unpublished (1997)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs-r@mail.nih.gov  
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Seq primer: M13 Forward  
 POLYA=Yes.

## FEATURES

source

Location/Qualifiers  
 1..302  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2717499"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP Sub3"  
 /note="Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Site1: Not I; Site2: Eco RI; The NCI CGAP Sub3 library is a subtracted library derived from the NCI CGAP Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalized or subtracted NCI CGAP libraries:  
 NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP Pr38, NCI CGAP Co10, NCI CGAP Co16, NCI CGAP Kids, NCI CGAP Kid12, NCI CGAP Kid3, NCI CGAP Kid5, NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP Co8, NCI CGAP CLL1, NCI CGAP Lei2, NCI CGAP Brn23, NCI CGAP Lu5, NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP GC4, NCI CGAP GC6,  
 NCI CGAP Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:  
 NCI CGAP Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs 132376-132391), 1456008-1456775, 1500552-1502855); NCI CGAP Kids pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI CGAP Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439); NCI CGAP GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI CGAP Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759, 1101192-1101959, 1217928-1220615); NCI CGAP Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255, 114584-1145351). Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.  
 TAG\_TISSUE=colon  
 TAG\_LIB=NCI CGAP\_Co10  
 TAG\_SEQ=AAACG"

## ORIGIN

Query Match 100.0%; Score 22; DB 10; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  
 Db

## RESULT 7

T74926/c

## LOCUS

DEFINITION

YC58b07.r1 Stratagene liver (#937224) Homo sapiens cDNA clone

IMAGE:84853 5', mRNA sequence.

T74926

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 344)

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissee, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

PUBMED

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 279

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

Seq primer: M13RP1

High quality sequence stop: 279.

Location/Qualifiers

1..344

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:501910"

/db\_xref="taxon:9606"

/clone="IMAGE:84853"

/sex="male"

/dev\_stage="49 years old"

/lab\_host="SOLR cells (kanamycin resistant)"

/clone\_lib="Stratagene liver (#937224)"

/note="Organ: liver; Vector: pBluescript SK; Site.1:

EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Hepatectomy from normal male caucasian. Average

insert size: 1.1 kb; Uni-ZAP XR Vector: -5' adaptor

sequence: 5' GAATTCGCACGAG 3' -3' adaptor sequence: 5'

CTCAGTGTTCAGGTAGAAAAGCC 16

CTCAGTGTTCAGGTAGAAAAGCC 16

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

0; Mismatches

0; Indels

0; Gaps

0;

OY

Db

## RESULT 8

AA448887

## LOCUS

DEFINITION

zx11f09.sl Soares\_total\_fetus\_Nb2HP8\_9w Homo sapiens cDNA clone

345 bp mRNA linear EST 04-JUN-1997

IMAGE:786185 3', mRNA sequence.

AA448887  
VERSION AA448887.1 GI:2162557  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
WashU-Merck EST Project 1997  
Unpublished (1997)

TITLE Unpublished (1997)  
JOURNAL  
COMMENT

CONTACT: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -41ml3 fwd. Et from Amersham  
High quality sequence stop: 299.

FEATURES  
source  
1. .345  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:5983804"  
/db\_xref="taxon:9606"  
/clone="IMAGE:786185"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares total\_fetus Nb2HF8\_9w"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCATCTGAGTGGAGCGCGCTTAATTTTCTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN  
Query Match 100.0%; Score 22; DB 9; Length 345;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCATGTTACAGTGAAGAACCC 22  
|||||  
Db 46 GCATGTTACAGTGAAGAACCC 67

RESULT 9  
LOCUS R00504  
DEFINITION ye74c07.s1 Soares fetal liver spleen INFLS Homo sapiens CDNA clone  
IMAGE:123468 3', mRNA sequence.  
R00504  
VERSION R00504.1 GI:750240  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and

Wilson,R.  
The WashU-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

Insert Size: 599  
High quality sequence stops: 321 Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 599 Std Error: 0.00  
Seq primer: -21ml3  
High quality sequence stop: 321.

FEATURES  
Location/Qualifiers  
1. .420  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:476013"  
/db\_xref="taxon:9606"  
/clone="IMAGE:123468"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGATTAATAAGATCTTTTCTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN  
Query Match 100.0%; Score 22; DB 14; Length 420;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCATGTTACAGTGAAGAACCC 22  
|||||  
Db 64 GCATGTTACAGTGAAGAACCC 85

RESULT 10  
LOCUS AA228376  
DEFINITION nc39e09.s1 NCI\_CGAP\_Pr2 Homo sapiens CDNA clone IMAGE:1010536, mRNA  
sequence.  
AA228376  
VERSION AA228376.1 GI:1849927  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 425)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuqui,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html  
Seq primer: -41m13 fwd. RT from Amersham  
High quality sequence stop: 403.

## FEATURES

source

1. .425  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1010536"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr2"  
/note="Vector: pAMP10; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

## ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 425;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAGCC 22  
|||||  
Db 80 GCATGTTACAGGTAGAAAGCC 101

RESULT 11  
AW571588 449 bp mRNA linear EST 13-MAR-2000  
LOCUS xx33e03.x1 NCI CGAP\_Utl1 Homo sapiens cDNA clone IMAGE:2839324 3',  
DEFINITION mRNA sequence.  
ACCESSION AW571588  
VERSION AW571588.1 GI:7236319  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 449)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/iresources.shtml

Possible reversed clone: polyT not found  
Seq primer: -40UP from Gibco  
High quality sequence stop: 367.

## FEATURES

source

1. .449  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2839324"

/tissue\_type="well-differentiated endometrial adenocarcinoma, 7 pooled tumors"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP\_Utl1"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"

## ORIGIN

Query Match 100.0%; Score 22; DB 10; Length 449;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAGCC 22  
|||||  
Db 426 GCATGTTACAGGTAGAAAGCC 447

## RESULT 12

AA004521 457 bp mRNA linear EST 07-MAY-1997  
LOCUS zh89h01.s1 Soares\_fetal\_liver\_spleen\_INFLS\_S1 Homo sapiens cDNA  
DEFINITION clone IMAGE:428497 3', mRNA sequence.  
ACCESSION AA004521  
VERSION AA004521.1 GI:1448098  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 457)  
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Pavello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Marais, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R., and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)

97044478

8889549

## COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1202 Std Error: 0.00

Seq primer: mob.REGA+EF.

Location/Qualifiers

## FEATURES

source

1. .457  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:1328266"  
/db\_xref="taxon:9606"  
/clone="IMAGE:428497"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares\_fetal\_liver\_spleen\_INFLS\_S1"  
/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; This is a subcloned version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5',  
AACTGGAAGATTAAATTAAGACTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified p7T3 vector. Library



went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 457;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
Db 44 GCATGTTACAGGTAGAAAAGCC 65

## RESULT 13

AW966858 505 bp mRNA linear EST 01-JUN-2000  
DEFINITION EST378932 MAGE resequences, MAGU Homo sapiens cDNA, mRNA sequence.

AW966858  
VERSION AW966858.1 GI:8156694

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and Quackenbush, J.

TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 13,200 element cDNA microarray

## JOURNAL

COMMENT Unpublished (2000)

Contact: John Quackenbush

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 3528

Fax: 301 838 0208

Email: johnq@tigr.org

Plate: 234

Seq primer: Forward.

Location/Qualifiers

1..505

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_lib="MAGE resequences, MAGJ"

/note="Vector: pBluescriptSKm"

## ORIGIN

Query Match 100.0%; Score 22; DB 10; Length 505;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22  
|||||  
Db 468 GCATGTTACAGGTAGAAAAGCC 489

## RESULT 14

AA228367 512 bp mRNA linear EST 21-AUG-1997  
DEFINITION nc39d08.s1 NCI\_CGAP\_Pr2 Homo sapiens cDNA clone IMAGE:1010511, mRNA sequence.

AA228367

VERSION AA228367.1 GI:1849918

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

## JOURNAL

## COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,

M.D., Michael Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.

cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -41m13 fwd. ET from Amersham

High quality sequence stop: 451.

## FEATURES

## source

1..512

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1010511"

/sex="Male"

/dev\_stage="45 years old"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_Pr2"

/note="Vector: PAMP10; Site 1: NotI; Site 2: EcoRI; 1st

strand cDNA was primed with oligo(dT)17 on 50 ng of

DNase-treated, total cellular RNA obtained from

5,000-10,000 microdissected preneoplastic cells

histologically-determined to be prostatic intraepithelial

neoplasia 2 (PIN2) cells. Double-stranded cDNA was

ligated to EcoRI adaptors, 5 cycles of PCR applied to the

cDNA with an adaptor-specific primer, and the resulting

PCR product subcloned into pAMP10 by the UDG-cloning

method (Life Technologies). Average insert size is 600

bp. NOTE: Not directionally cloned. This library was

constructed by David Krizman."

## ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 512;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATGTTACAGGTAGAAAAGCC 22

|||||

Db 80 GCATGTTACAGGTAGAAAAGCC 101

## RESULT 15

## LOCUS

## DEFINITION

AW247896 522 bp mRNA linear EST 07-JAN-2000  
2820683.3prime NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:2820683 3', mRNA sequence.

AW247896

VERSION AW247896.1 GI:6590889

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

## AUTHORS

NIH-MGC http://mgi.nci.nih.gov/.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

## JOURNAL

## COMMENT

Other ESTs: 2820683.5prime

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: DCTD/DTG cDNA Library Preparation: Ling

Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.

Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing

project Clone distribution: MGC clone distribution information can

be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality

Scores: PHRED from University of Washington Genome Center. Vector

Trimming: cross\_match from University of Washington Genome Center

PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley  
Drosophila Genome Project. University of Washington Genome Center:  
<http://www.genome.washington.edu> Polyadenylation: Based upon the  
presence of a xhoI site followed by a run of 14 or more T residues  
at the beginning of the sequence, this cDNA insert was  
polyadenylated.

Plate: L16M4 row: N column: 12  
High quality sequence stop: 402.

FEATURES  
source

Location/Qualifiers  
1..522  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2820683"  
/tissue\_type="small cell carcinoma"  
/cell\_line="MGC3"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 7"

/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 100.0%; Score 22; DB 10; Length 522;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCATGTTACAGTAGAAAGCC 22  
|||  
Db 485 GCATGTTACAGTAGAAAGCC 506

Search completed: April 10, 2004, 19:48:50  
Job time : 78.2102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 11:11:52 ; Search time 84.7999 Seconds  
(without alignments)  
10733.545 Million cell updates/sec

Title: US-09-743-825-8

Perfect score: 21

Sequence: 1 ctggcgatctgaagagctctg 21

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	17.8	84.8	1001	9	HSM5HV1	AF070072 Homo sapi
C 2	17.8	84.8	2455	6	AX822172	AX822172 Sequence
C 3	17.8	84.8	2455	6	AX825812	AX825812 Sequence
C 4	17.8	84.8	6820	9	HSM5HV1	AF049988 Homo sapi
C 5	17.8	84.8	28772	9	HSA012008	AJ012008 Homo sapi
C 6	17.8	84.8	53125	9	AL844216	AL844216 Human DNA
C 7	17.8	84.8	54146	9	BX248244	BX248244 Human DNA
C 8	17.8	84.8	74769	5	BX649467	BX649467 Zebrafish
C 9	17.8	84.8	100000	9	AP000503S2	AP000504 Homo sapi
C 10	17.8	84.8	125350	2	AC020768	AC020768 Homo sapi
C 11	17.8	84.8	136493	9	AL662899	AL662899 Human DNA
C 12	17.8	84.8	152027	2	AC132297	AC132297 Mus muscu
C 13	17.8	84.8	163682	9	HSDJ71117	AL132713 Human DNA
C 14	17.8	84.8	179894	9	AL662834	AL662834 Human DNA
C 15	17.8	84.8	180283	9	AF134726	AF134726 Homo sapi
C 16	17.8	84.8	184666	9	DJ201G34	AF129756 Homo sapi
C 17	17.8	84.8	227968	10	AC098723	AC098723 Mus muscu
C 18	17.8	84.8	237515	2	AC126059	AC126059 Rattus no
C 19	17.8	84.8	247239	2	AC091376	AC091376 Rattus no
C 20	17.8	84.8	248441	2	AC098301	AC098301 Rattus no
C 21	17.8	84.8	252216	2	AC128970	AC128970 Rattus no
C 22	17.8	84.8	256608	10	AL589699	AL589699 Mouse DNA
C 23	17.8	82.9	463	3	TMU24237	U24237 Tenebrio mo
C 24	17.4	82.9	34266	3	LMFL391	AL1135941 Leishmani
C 25	17.4	82.9	121995	2	AC091512	AC091512 Leishmani
C 26	17.4	82.9	242527	2	AC115405	AC115405 Rattus no
C 27	17.4	82.9	252053	2	AC097148	AC097148 Rattus no
C 28	17	81.0	137384	10	AC121893	AC121893 Mus muscu
C 29	17	81.0	186437	2	AC131918	AC131918 Mus muscu
C 30	17	81.0	285836	2	AC128556	AC128556 Rattus no
C 31	17	81.0	314250	1	BX294149	BX294149 Pirellula
C 32	16.8	80.0	300	9	HS170B6R	Z57285 H.sapiens C
C 33	16.8	80.0	1968	1	AF244992	AF244992 Vibrio ch
C 34	16.8	80.0	2447	5	BC049474	BC049474 Danio rer
C 35	16.8	80.0	10192	1	AE004222	AE004222 Vibrio ch
C 36	16.8	80.0	11108	1	AE012151	AE012151 Xanthomon
C 37	16.8	80.0	16294	1	AE000732	AE000732 Aquifex a
C 38	16.8	80.0	23775	1	BACYP1A	I47709 Bacillus su
C 39	16.8	80.0	40338	2	AC147414	AC147414 Homo sapi
C 40	16.8	80.0	56800	2	AC132198	AC132198 Homo sapi
C 41	16.8	80.0	66253	2	AC023372	AC023372 Homo sapi
C 42	16.8	80.0	69762	2	AC137727	AC137727 Homo sapi
C 43	16.8	80.0	76527	9	AC093459	AC093459 Homo sapi
C 44	16.8	80.0	98340	9	AL353716	AL353716 Human DNA
C 45	16.8	80.0	103186	9	AL158218	AL158218 Human DNA

ALIGNMENTS

RESULT 1  
HSM5HV1/c  
LOCUS  
DEFINITION Homo sapiens Muts homolog (MSH5) gene, exons 1 and 2.  
ACCESSION AF070072  
VERSION AF070072.1 GI:3831706  
KEYWORDS  
SEGMENT  
SOURCE 1 of 8  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1001)  
AUTHORS Winand,N.J., Panzer,J.A. and Kolodner,R.D.  
TITLE Cloning and characterization of the human and Caenorhabditis



exon 766. .958  
/gene="MSH5"  
/number=1  
exon 1192. .1351  
/gene="MSH5"  
/number=2  
exon 1901. .2024  
/gene="MSH5"  
/number=3  
exon 3604. .3684  
/gene="MSH5"  
/number=4  
exon 3835. .3897  
/gene="MSH5"  
/number=5  
exon 4631. .4752  
/gene="MSH5"  
/number=6  
exon 4917. .5026  
/gene="MSH5"  
/number=7  
exon 5273. .5308  
/gene="MSH5"  
/number=8  
exon 5964. .6046  
/gene="MSH5"  
/number=9

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 6820;  
Best Local Similarity 90.5%; Pred. No. 1.8e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGAAGAGTCTG 21  
Db 625 CTGGCGTATCTGAAGAGTCTG 605

RESULT 5  
HSA012008 28772 bp DNA linear PRI 12-JUL-1999  
LOCUS Homo sapiens genes encoding RNCC protein, DDAH protein, Ly6-C protein, Ly6-D protein and immunoglobulin receptor.  
DEFINITION  
ACCESSION AJ012008.1 GI:5304874  
VERSION  
KEYWORDS DDAH protein; dimethyl arginine dimethylaminohydrolase; G6 gene; G6A gene; G6B gene; G6C gene; G6D gene; G7 gene; Ig-R; immunoglobulin receptor; Ly6-C protein; Ly6-D protein; regulatory nuclear chlorine ion channel protein; RNCC protein.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Ribas,G., Neville,M., Wixon,J.L., Cheng,J. and Campbell,R.D.  
TITLE Genes encoding three new members of the leukocyte antigen 6 superfamily and a novel member of Ig superfamily, together with genes encoding the regulatory nuclear chloride ion channel protein (hRNCC) and an N omega-N omega-dimethylarginine dimethylaminohydrolase homologue, are found in a 30-kb segment of the MHC class III region  
J. Immunol. 163 (1), 278-287 (1999)  
MEDLINE 99316314  
PUBMED 10384126  
REFERENCE 2 (bases 1 to 28772)  
AUTHORS Ribas,G.  
TITLE Direct Submission  
JOURNAL Submitted (22-JUN-1998) Ribas G., Biochemistry, MRC Immunochemistry Unit, South Parks, Oxford OX1 3QU, UK  
FEATURES  
source 1. .28772  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"

gene /db\_xref="taxon:9606"  
complement(143. .1393)  
/gene="G7"  
exon complement(143. .250)  
/gene="G7"  
/number=3  
repeat\_region complement(305. .414)  
/rpt\_family="SINE/MIR"  
repeat\_region complement(431. .734)  
/rpt\_family="Alu Sx"  
exon complement(800. .959)  
/gene="G7"  
/number=2  
exon complement(1351. .1393)  
/gene="G7"  
/number=1  
repeat\_region 1826. .1895  
/rpt\_family="MIR"  
repeat\_region complement(2159. .2252)  
/rpt\_family="LINE/L2"  
repeat\_region 2255. .2309  
/standard\_name="9N-1"  
/rpt\_family="(AC)n"  
gene 2355. .10832  
/gene="G6"  
promoter 2355. .2394  
/gene="G6"  
repeat\_region complement(3980. .4280)  
/rpt\_family="Alu Sq"  
mRNA join(4904. .5160,7156. .7265,7458. .7583,7747. .7853,  
9020. .9201,10415. .10832)  
/gene="G6"  
/evidence=experimental  
join(5122. .5160,7156. .7265,7458. .7583,7747. .7853,  
9020. .9201,10415. .10576)  
/gene="G6"  
/standard\_name="regulatory nuclear chlorine ion channel protein"  
/codon\_start=1  
/evidence=experimental  
/product="RNCC protein"  
/protein\_id="CAB46078.1"  
/db\_xref="GI:5304875"  
/db\_xref="REMTREMBL:CAB46078"  
translation="MAREQPOVELFVKAGSDGKIGNCPFSORLPFWLWLGKVTENV  
TVDKRTETVQKLCPCGSQLPFLLYGTEVHTDTNKLEFLAVLCPPRYPKLAALNPE  
SNTAGLDIFAKFSAYIKSNFALNDNLEKLLKALKVDNYLTSLPESVDTSAEDE  
GVQRKFLDGNELTLADCNLPKHLVQVCKYRGFTIPEAFRGVHRYLSNAYAREE  
FASTCPDDEIEELAYEQVAKALK"  
5572. .5612  
/standard\_name="82.1I"  
repeat\_region /rpt\_family="(AC)n"  
8114. .8278  
/rpt\_family="MIR"  
repeat\_region complement(8282. .8567)  
/rpt\_family="Alu Sx"  
repeat\_region 8610. .8853  
/rpt\_family="Alu"  
repeat\_region complement(9678. .9816)  
/rpt\_family="LINE/L2"  
repeat\_region 9817. .10108  
/rpt\_family="Alu Sx"  
repeat\_region complement(10109. .10133)  
/rpt\_family="LINE/L2"  
gene 10841. .14379  
/gene="G6A"  
promoter 10841. .10880  
/gene="G6A"  
mRNA join(11156. .11367,12194. .12554,12674. .12773,12895. .12968,  
13127. .13246,13727. .13876,14066. .14379)  
/gene="G6A"  
/product="transcript A"

```

mRNA
join(11635..11736,12198..12554,12674..12773,12895..12968,
13127..13246,13727..13876,14066..14379)
/gene="G6A"
/repeat_region complement(21882..21928)
/rept_family="Alu Sc"
gene
join(12052..12554,12674..12773,12895..12968,13127..13246,
13727..13876,14066..14379)
/gene="G6A"
/rept_family="MIR"
CDS
join(12258..12554,12674..12773,12895..12968,13127..13246,
13727..13876,14066..14182)
/gene="G6A"
/codon_start=1
/standard_names="dimethyl arginine dimethylaminohydrolase"
/evidence=experimental
/product="DDAH protein"
/protein_id="CAB46079.1"
/db_xref="GI:5304876"
/db_xref="GOA:O95865"
/db_xref="SWISS-PROT:O95865"
/translation="MGTFGEGLGRCSHALIRGVPESLASGEGAGLPAALDAKQRE
HGVGGKQRIGLQLLPEESLPLGLDGTAVIOGDTALITRPMSPARRBEVDG
VRKALQDGLRIVEIGDENATLDGTVLFTGREFVGLSKWTHRGABIVADTFRDEA
VSTVPSGSHLRCGCGGPRVTVAGSSDAQAQAVRAMAVLTDHPYASLTLPDAAA
DCLFLRPLGVPVPPLLHRRGGDLFNSQEQALQLSDVTLVPVSCSELEKAGAGLSLC
LVLSRPHS"
complement(14899..15119)
/rept_family="LINE/L2"
/repeat_region complement(15199..15759)
/rept_family="LINE/L2"
repeat_region
15793..16088
/rept_family="Alu Y"
complement(16345..18260)
/gene="G6B"
CDS
complement(join(16345..16449,16575..16654,16792..16832,
17187..17277,17433..17780,17975..18035))
/gene="G6B"
/feature="first and last exon not identified"
/codon_start=1
/evidence=not_experimental
/product="immunoglobulin receptor"
/protein_id="CAB46080.1"
/db_xref="GI:5304877"
/db_xref="REMTREMBL:CAB46080"
/translation="WAVFLQLPLLSSRAQGNPGASLDGRPGDRVNLSCGVSPIRW
WVAPSPACKGSKGRPILWASSGTETVPLQFVGRKLSDSGRRLLELLSAGD
SGTFPCRHEDESSTVJLVLDRTYKAPGTHGVSVPQLLIPLLGAGLVIGLALG
LVWMLHRLPQPTIRPLPRFAPLVKTEPQRPVKEEPKIPGDLQPEPSLLYADLHLA
LSRRLSLTADPADASTTYAVVV"
complement(18221..18260)
/gene="G6B"
promoter
18289..18426
/rept_family="Alu Jb"
repeat_region complement(18557..18841)
/rept_family="Alu Sc"
gene
19257..22767
/gene="G6C"
promoter
19257..19296
/gene="G6C"
mRNA
join(19683..19788,21212..21322,22105..22767)
/gene="G6C"
CDS
join(19737..19788,21212..21322,22105..22319)
/gene="G6C"
/codon_start=1
/product="Ly6-C protein"
/protein_id="CAB46081.1"
/db_xref="GI:5304878"
/db_xref="REMTREMBL:CAB46081"
/translation="MKALMLTLVLLCWVGNADIRCHSKYKVPVLGCVDROSCLREP
QQCLTHAYLGNMWFVSNLRCGTPEPCQEAFFNQTRKXGLGTYNTCCNKDNCNSAGP
RPTPALGLVFLTSLAGLGLWLLH"
complement(21459..21543)
/rept_family="SINE/MIR"
repeat_region
21579..21874

```

```

/rept_family="Alu Sc"
complement(21882..21928)
/rept_family="MIR"
/gene="G6D"
complement(join(23611..23834,25789..25911,26003..26057))
/gene="G6D"
/codon_start=1
/evidence=experimental
/product="Ly6-D protein"
/protein_id="CAB46082.1"
/db_xref="GI:5304879"
/db_xref="REMTREMBL:CAB46082"
/translation="MKQFVGLISLLGALGNRMRCYNGCGSPSSSCKEAVTTGGE
GRPQPGDEQKLPQNPFTVLIHQHPACVAAHHCNQVETESVGDVTYTPAHRDCYLGDL
NSAVASHVAPAGILAAAATLTCLPLGLWSG"
23960..24124
/rept_family="MER42"
complement(24150..24447)
/rept_family="Alu Sp"
repeat_region 24448..24475
Query Match 84.8% Score 17.8; DB 9; Length 28772;
Best Local Similarity 90.5%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CTGCGGTATCTGAAGAGTCTG 21
Db 1526 CTGCGGTAGTGAAGAGTCTG 1546
RESULT 6
AL844216/c
LOCUS Human DNA sequence from clone XXbac-350E15 on chromosome 6,
DEFINITION complete sequence.
ACCESSION AL844216
VERSION AL844216.3 GI:21955728
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 53125)
AUTHORS Tracey, A.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 25, 2002 this sequence version replaced gi:21953192.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
database can be found at

```

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6/MHC>.  
Xxbac-350E15 is from a DNA-arts QBL human bac library VECTOR: pBelOBAC11.

#### FEATURES

source  
1. .53125  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="Xxbac-350E15"  
/clone\_lib="DNA-arts-BAC.1-QBL.1"

#### ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 53125;  
Best Local Similarity 90.5%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGAAGAGTCTG 21  
Db 31251 CTGCCGTAGCTGAAGAGTCTG 31231

#### RESULT 7

EX248244/c  
LOCUS  
DEFINITION  
Human DNA sequence from clone DASS-312M4 on chromosome 6, complete sequence.  
ACCESSION  
BX248244  
VERSION  
BX248244.7 GI:33636283  
KEYWORDS  
HTG.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 54146)  
Wood, J.  
Direct Submission  
Submitted (13-AUG-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
On Aug 13, 2003 this sequence version replaced gi:33620366.

#### COMMENT

Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: humquery@sanger.ac.uk  
-----  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WormPEP; Information on the WormPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep). This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6/MHC>  
DASS-312M4 is from a DNA-arts SSTO human bac library VECTOR: pBelOBAC11.

#### FEATURES

source  
1. .54146  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="DASS-312M4"  
/clone\_lib="DNA-arts-BAC.1-SSTO.1"

#### ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 54146;  
Best Local Similarity 90.5%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGAAGAGTCTG 21  
Db 52444 CTGCCGTAGCTGAAGAGTCTG 52424

#### RESULT 8

EX649467/c  
LOCUS  
DEFINITION  
Zebrafish DNA sequence from clone CH211-117L16 in linkage group 4, complete sequence.  
ACCESSION  
EX649467  
VERSION  
EX649467.5 GI:37699925  
KEYWORDS  
HTG.  
SOURCE  
Danio rerio (zebrafish)  
ORGANISM  
Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
REFERENCE  
1 (bases 1 to 74769)  
Phillimore, B.  
Direct Submission  
Submitted (16-OCT-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk  
On Oct 16, 2003 this sequence version replaced gi:37651136.

#### COMMENT

Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: zfish-help@sanger.ac.uk  
-----  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WormPEP; Information on the WormPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep). Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.  
Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those

beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, Washu). For further information see [http://www.sanger.ac.uk/Projects/D\\_rerio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml)  
CH211-117116 is from a CHORI-211 BAC library  
VECTOR: pTARBAC2.1.

FEATURES  
source  
1..74769  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="CH211-117116"  
/clone\_lib="CHORI-211"

# ORIGIN

Query Match 84.8%; Score 17.8; DB 5; Length 74769;  
Best Local Similarity 90.5%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
||||| ||||| ||||| |||||

Db 59345 CTGGCGGATCTGAAGAGTCTG 59325

## RESULT 9

AP00050382

LOCUS

DEFINITION

AP00050382 Homo sapiens genomic DNA, chromosome 6p21.3, HLA Class I region, section 3/20.

ACCESSION

AP000504 BA000025

VERSION

AP000504.1 GI:5926691

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Shiina,S., Tamiya,G., Oka,A. and Inoko,H.

Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region

Published Only in Database (1999)

2 (bases 1 to 100000)

Hirakawa,M., Yamaguchi,H., Imai,K. and Shimada,J.

Direct Submission

Submitted (21-SEP-1999) Mika Hirakawa, Japan Science and Technology

Corporation (JST), Advanced Databases Department: 5-3, Yonbancho,

Chiyoda-ku, Tokyo 102-0081, Japan (E-mail:mika@tokyo.jst.go.jp,

URL:htp://www-alis.tokyo.jst.go.jp/, Tel:81-3-5214-8491,

Fax:81-3-5214-8470)

This sequence is conducted by Tokai University as a JST sequencing

Team.

Principal Investigator: Hidetoshi Inoko Ph.D

Phone:+81-463-93-1121, Fax:+81-463-94-8884,

The sequence is submitted by Human Genome Sequencing in ALIS

project of JST

Japan Science and Technology Corporation (JST)

5-3, Yonbancho, Chiyoda-ku, Tokyo, 102-0081 Japan

For further information about this sequence, please visit our

sequence archive Web site (<http://www-alis.tokyo.jst.go.jp/HGS/top.html>) or send email to [webmaster@www-alis.tokyo.jst.go.jp](mailto:webmaster@www-alis.tokyo.jst.go.jp).

Location/Qualifiers

1..100000

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosomes="6"

/map="6p21.3"

# ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 100000;

Best Local Similarity 90.5%; Pred. No. 1.5e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
||||| ||||| ||||| |||||

Db 4151 CTGGCGTAGTGAAGAGTCTG 4171

## RESULT 10

AC020768

LOCUS

DEFINITION

AC020768 Homo sapiens clone RP11-11L21, WORKING DRAFT SEQUENCE, 24 unordered

pieces.

AC020768

AC020768.3 GI:7341894

VERSION

HTG; HTGS PHASE1; HTGS\_DRAFT.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 125350)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens, clone RP11-11L21

Unpublished

2 (bases 1 to 125350)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,

Boguslavsky,I., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,

Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,

Deaillano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,

Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,

Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Johnson,R., Jones,C., Kam,L., Karatas,A., Klein,J.,

Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,

Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,

McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,

Norman,C.H., O'Connor,T., O'Donnell,P., Olivart,T.M., Peterson,K.,

Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,

Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,

Zimmer,A. and Zody,M.

Direct Submission

Submitted (09-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 30, 2000 this sequence version replaced GI:6778558.

All repeats were identified using RepeatMasker:

Smit,A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L3221

Center clone name: 11\_L121

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 114079 bases at least Q40

Consensus quality: 119463 bases at least Q30

Consensus quality: 121683 bases at least Q20

Insert size: 133000; agarose-fp

Insert size: 123050; sum-of-contigs

Quality coverage: 4.3 in Q20 bases; agarose-fp

Quality coverage: 4.7 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 24 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 1570: contig of 1570 bp in length



```
* 1571 1670: gap of 100 bp
* 1671 3378: contig of 1708 bp in length
* 3379 3478: gap of 100 bp
* 3479 4736: contig of 1258 bp in length
* 4737 4836: gap of 100 bp
* 4837 6654: contig of 1818 bp in length
* 6655 6754: gap of 100 bp
* 6755 8445: contig of 1691 bp in length
* 8446 8545: gap of 100 bp
* 8546 10525: contig of 1980 bp in length
* 10526 10625: gap of 100 bp
* 10626 12550: contig of 1925 bp in length
* 12551 12650: gap of 100 bp
* 12651 16081: contig of 3431 bp in length
* 16082 16181: gap of 100 bp
* 16182 19200: contig of 3019 bp in length
* 19201 19300: gap of 100 bp
* 19301 21130: contig of 1830 bp in length
* 21131 21230: gap of 100 bp
* 21231 23618: contig of 2388 bp in length
* 23619 23718: gap of 100 bp
* 23719 26777: contig of 3059 bp in length
* 26778 26877: gap of 100 bp
* 26878 29428: contig of 2551 bp in length
* 29429 32735: contig of 3207 bp in length
* 32736 32835: gap of 100 bp
* 32836 37811: contig of 4976 bp in length
* 37812 37911: gap of 100 bp
* 37912 42611: contig of 4700 bp in length
* 42612 42711: gap of 100 bp
* 42712 49022: contig of 6311 bp in length
* 49023 49122: gap of 100 bp
* 49123 56304: contig of 7182 bp in length
* 56305 56404: gap of 100 bp
* 56405 63059: contig of 6655 bp in length
* 63060 63159: gap of 100 bp
* 63160 70856: contig of 7697 bp in length
* 70857 70955: gap of 100 bp
* 70956 79641: contig of 8685 bp in length
* 79642 79741: gap of 100 bp
* 79742 90952: contig of 11211 bp in length
* 90953 91052: gap of 100 bp
* 91053 100495: contig of 9443 bp in length
* 100496 100595: gap of 100 bp
* 100596 125350: contig of 24755 bp in length.
```

## FEATURES

```
source      Location/Qualifiers
            1..125350
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /clone="RP11-11L21"
             /clone_lib="RPC1-11 Human Male BAC"
            misc_feature 1..1570
            misc_feature 1671..3378
            misc_feature /note="assembly_fragment"
            misc_feature /note="assembly_fragment"
            misc_feature 3479..4736
            misc_feature /note="assembly_fragment"
            misc_feature 4837..6654
            misc_feature /note="assembly_fragment"
            misc_feature 6755..8445
            misc_feature /note="assembly_fragment"
            misc_feature 8546..10525
            misc_feature /note="assembly_fragment"
            misc_feature 10626..12550
            misc_feature /note="assembly_fragment"
            misc_feature 12651..16081
            misc_feature /note="assembly_fragment"
            misc_feature 16182..19200
            misc_feature /note="assembly_fragment"
            misc_feature 19301..21130
            misc_feature /note="assembly_fragment"

Query Match      84.8%; Score 17.8; DB 2; Length 125350;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CTGGCGTATCTGAAGAGTCTG 21
Db      68860  CTGCCGTAGCTGAAGAGTCTG 68880

RESULT 11
AL662899/c
LOCUS   AL662899 Human DNA sequence from clone Xbac-32j3 on chromosome 6, complete
DEFINITION Human DNA sequence from clone Xbac-32j3 on chromosome 6, complete
sequence.
ACCESSION AL662899
VERSION   AL662899.5 GI:18650693
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1
  Tracey, A.
  Direct Submission
  Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  humquery@sanger.ac.uk
  On Feb 11, 2002 this sequence version replaced gi:18477418.
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the
  variation annotation may not be found in the sequence submission
  corresponding to the overlapping clone, as we submit sequences with
  only a small overlap as described above.
  This sequence was finished as follows unless otherwise noted: all
  regions were either double-stranded or sequenced with an alternate
  chemistry or covered by high quality data (i.e., phred quality >=
  30); an attempt was made to resolve all sequencing problems, such
  as compressions and repeats; all regions were covered by at least
```



[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>

RP1-71117 is from the library RPI-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pCYPAC2

This sequence is the entire insert of clone RP1-71117.

FEATURES  
 source  
 Location/Qualifiers  
 1..163682  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosomes="6"  
 /map="p21.2-22.1"  
 /clone="RP1-71117"  
 /clone\_lib="RPI-1"

## ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 163682;  
 Best Local Similarity 90.5%; Pred. No. 1.4e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
 |||||  
 Db 71862 CTGGCGTAGCTGAAGAGTCTG 71862

## RESULT 14

AL662834/c  
 LOCUS  
 DEFINITION  
 Human DNA sequence from clone Xbac-40G17 on chromosome 6, complete sequence.

ACCESSION  
 AL662834  
 VERSION  
 AL662834.8 GI:18673924  
 KEYWORDS  
 HTG.

SOURCE  
 Homo sapiens (human)

ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS  
 Tracey A.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)

COMMENT  
 humquery@sanger.ac.uk Clone requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)  
 On Feb 14, 2002 this sequence version replaced gi:18491411.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; SW., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)

from a CHORI-502 human bac - COX cell line library VECTOR: PTARBAC2.1

This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6/MHC>.

## FEATURES

source  
 Location/Qualifiers  
 1..179894  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosomes="6"  
 /clone="XXbac-40G17"  
 /clone\_lib="CHORI-502"

## ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 179894;  
 Best Local Similarity 90.5%; Pred. No. 1.4e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
 |||||  
 Db 3801 CTGGCGTAGCTGAAGAGTCTG 3781

## RESULT 15

AF134726  
 LOCUS  
 DEFINITION  
 Homo sapiens BAC clone 215012 NG35, NG36, G9A, NG22, G9, HSP70-2, HSP70-1, HSP70-HOM, snRNP, G7A, NG37, NG23, and MutSH5 genes, complete cds.

ACCESSION  
 AF134726  
 VERSION  
 AF134726.1 GI:4529886  
 KEYWORDS

SOURCE  
 Homo sapiens (human)

ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 1 (bases 1 to 180283)  
 AUTHORS  
 Rowen, L., Qin, S., Madan, A., Dickhoff, R., Dors, M., Madan, A., Hicks, P., Loretz, C., Ratcliffe, A., Abbasi, N., Shaffer, T. and Hood, L.

TITLE  
 Sequence of the human major histocompatibility complex class III region

JOURNAL  
 Unpublished

REFERENCE  
 2 (bases 1 to 180283)  
 AUTHORS  
 Rowen, L.

TITLE  
 Direct Submission

JOURNAL  
 Submitted (15-MAR-1999) Department of Molecular Biotechnology, Box 357730 University of Washington, Seattle, WA 98195, USA

COMMENT  
 Sequencing methodology: high redundancy shotgun in plasmids. Interspersed repeats were identified with RepeatMasker (available from <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). This sequence overlaps cosmid S22A (AF019413) by 2363 bases and BAC 210G24 (AF129756) by 12177 bases.

## FEATURES

source  
 Location/Qualifiers  
 1..180283  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosomes="6"  
 /maps="6p21.3"  
 /clone\_lib="RPC111"  
 /clone="BAC 215012"  
 /note="Major histocompatibility complex class III region"

misc\_feature  
 1..2363

repeat\_region  
 complement(103..402)

variation  
 /rpt\_family="AluSx"  
 /rpt\_type=dispersed  
 408..409

repeat\_region  
 /note="cosmid S22A: ttct; BAC 215012: tt"  
 /replace="ttct"  
 complement(420..480)

repeat\_region  
 /rpt\_family="L1PA16"  
 /rpt\_type=dispersed  
 complement(458..732)

repeat\_region  
 /rpt\_family="L1PA13"  
 /rpt\_type=dispersed

repeat_region	complement(834..1134)	/rpt_family="AluSx"	/rpt_type=dispersed	repeat_region	/rpt_family="AluSx"	/rpt_type=dispersed
variation	1134..1136	/note="cosmid S22A: ct; BAC 215012: ctt"			complement(7029..7330)	
repeat_region	/replace="ct"	complement(1139..1257)		repeat_region	/rpt_type=dispersed	
variation	1445..1446	/rpt_family="AluSg1"			7486..7688	
repeat_region	/note="cosmid S22A: ctt; BAC 215012: ct"	/rpt_type=dispersed		repeat_region	/rpt_family="AluJb"	
variation	1843	/replace="ctt"			7689..7851	
repeat_region	complement(1446..1746)	/rpt_family="AluSx"		repeat_region	/rpt_family="L1MB8"	
variation	2430..2489	/rpt_type=dispersed			7852..8150	
repeat_region	/replace="t"	complement(1446..1746)		repeat_region	/rpt_family="AluSg"	
variation	2430..2489	/rpt_type=dispersed			8151..8263	
repeat_region	/rpt_family="MLT1A2"	complement(2831..2884)		repeat_region	/rpt_family="L1MB6"	
variation	3813..4084	/rpt_type=dispersed			8290..8425	
repeat_region	/rpt_family="LINE2"	complement(3091..3392)		repeat_region	/rpt_family="FLAM_C"	
variation	4086..4375	/rpt_type=dispersed			8451..8730	
repeat_region	/rpt_family="AluSg"	complement(3719..3790)		repeat_region	/rpt_type=dispersed	
variation	4781..4910	/rpt_type=dispersed			8737..8839	
repeat_region	/rpt_family="FLAM_A"	complement(4417..4729)		repeat_region	/rpt_family="L1MB5"	
variation	5786..5873	/rpt_type=dispersed			8816..8972	
repeat_region	/rpt_family="AluSx"	complement(4939..4968)		repeat_region	/rpt_type=dispersed	
variation	5885..5938	/rpt_type=dispersed			9736..9882	
repeat_region	/rpt_family="MLT2E"	complement(4980..5112)		repeat_region	/rpt_family="FLAM_C"	
variation	6099..6389	/rpt_type=dispersed			9884..10185	
repeat_region	/rpt_family="AluJb"	complement(5167..5468)		repeat_region	/rpt_family="AluSg"	
variation	5955..6090	/rpt_type=dispersed			10189..10351	
repeat_region	/rpt_family="FLAM_C"	complement(5607..5738)		repeat_region	/rpt_family="AluSc"	
variation	6403..6446	/rpt_type=dispersed			10357..10419	
repeat_region	/rpt_family="LINE2"	complement(5786..5873)		repeat_region	/rpt_family="AluJb"	
variation	6099..6389	/rpt_type=dispersed			complement(10427..10711)	
repeat_region	/rpt_family="MER5A"	complement(5167..5468)		repeat_region	/rpt_family="AluJb"	
variation	5955..6090	/rpt_type=dispersed			complement(10871..11176)	
repeat_region	/rpt_family="MER5A"	complement(5607..5738)		repeat_region	/rpt_family="AluSg"	
variation	6403..6446	/rpt_type=dispersed			complement(12410..12541)	
repeat_region	/rpt_family="FLAM_C"	complement(5786..5873)		repeat_region	/rpt_family="FLAM_C"	
variation	6099..6389	/rpt_type=dispersed			complement(12577..12873)	
repeat_region	/rpt_family="MER5A"	complement(5167..5468)		repeat_region	/rpt_family="AluJo"	
variation	5955..6090	/rpt_type=dispersed			complement(13907..13948)	
repeat_region	/rpt_family="MER5A"	complement(5607..5738)		repeat_region	/rpt_type=dispersed	
variation	6403..6446	/rpt_type=dispersed			13413..13701	
repeat_region	/rpt_family="FLAM_C"	complement(5786..5873)		repeat_region	/rpt_family="AluSg"	
variation	6099..6389	/rpt_type=dispersed			complement(13702..13744)	
repeat_region	/rpt_family="MER5A"	complement(5167..5468)		repeat_region	/rpt_family="AT rich"	
variation	5955..6090	/rpt_type=dispersed			complement(13907..13948)	
repeat_region	/rpt_family="MER5A"	complement(5607..5738)		repeat_region	/rpt_family="MIR"	
variation	6403..6446	/rpt_type=dispersed			complement(13995..14046)	
repeat_region	/rpt_family="FLAM_C"	complement(5786..5873)		repeat_region	/rpt_type=dispersed	
variation	6099..6389	/rpt_type=dispersed			complement(13995..14046)	
repeat_region	/rpt_family="MER5A"	complement(5167..5468)		repeat_region	/rpt_family="MER57B"	
variation	5955..6090	/rpt_type=dispersed			complement(6711..7011)	

```
repeat_region complement(14378..14674)
/rpt_family="AluSx"
Query Match      84.8%; Score 17.8; DB 9; Length 180283;
Best Local Similarity 90.5%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CTGCGGTATCTGAAGAGTCTG 21
      ||||| ||||| ||||| |||||
Db 178287 CTGCGGTAGCTGAAGAGTCTG 178307
```

Search completed: April 10, 2004, 17:33:02  
Job time : 89.7999 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 14:24:52 ; Search time 8.80787 Seconds

(without alignments)  
10128.688 Million cell updates/sec

Title: US-09-743-825-8

Perfect score: 21

Sequence: 1 ctggcgatctgaagagctg 21

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	21	3	Aaz50445 EST R0050
2	21	100.0	25	8	ACI19942 Human mic
3	19.4	92.4	25	8	ACI19943 Human mic
4	17.8	84.8	2455	9	ADB54008 MSH5 geno
5	17.8	84.8	28772	6	ABK83555 Human CDN
6	16.4	78.1	2751	7	ACA52436 Prokaryot
7	16.4	78.1	3357	2	AAX20612 Polynucle
8	16.4	78.1	13435	4	AAI99253 Human exc
9	16.4	78.1	13435	4	AAI04187 Human rep
10	16.4	78.1	13435	5	AAI63603 Human kid
11	16.2	77.1	36	7	ACC69924 Human TRA
12	16.2	77.1	36	7	ACC69942 Human TRA
13	16.2	77.1	36	7	ACC69941 Human TRA
14	16.2	77.1	36	7	ACC69941 Human TRA
15	16.2	77.1	230	6	ABS69156 Novel mur
16	16.2	77.1	376	4	ABA26713 Probe #1
17	16.2	77.1	396	7	ABX05837 S. pneumo
18	16.2	77.1	411	7	ABX05835 S. pneumo
19	16.2	77.1	414	4	AAS55469 Streptoco
20	16.2	77.1	414	4	AAS55793 Streptoco
21	16.2	77.1	414	7	ACA49676 Prokaryot
22	16.2	77.1	439	4	AAS34770 cDNA enco
23	16.2	77.1	439	4	AAS31266 Human CDN

C 24	16.2	77.1	439	4	AAI63805	Aai63805 Human pol
C 25	16.2	77.1	439	4	AAS28081	Aas28081 Novel CDN
C 26	16.2	77.1	439	4	ABA06370	Abao6370 Human CDN
C 27	16.2	77.1	439	5	AAS29638	Aas29638 Human end
C 28	16.2	77.1	439	6	ABQ66590	Abq66590 Human pol
C 29	16.2	77.1	439	6	ABV83707	Abv83707 Human CDN
C 30	16.2	77.1	439	9	ADC10612	Adc10612 Human CDN
C 31	16.2	77.1	439	9	ADC45928	Adc45928 Human neo
C 32	16.2	77.1	597	2	AAT98628	Aat98628 DNA encod
C 33	16.2	77.1	600	2	AAX30765	Aax30765 Streptoco
C 34	16.2	77.1	850	4	AAH08228	Aah08228 Human CDN
C 35	16.2	77.1	1488	7	ADA71031	Ada71031 Rice gene
C 36	16.2	77.1	1587	7	ACA25275	ACA25275 Prokaryot
C 37	16.2	77.1	2540	4	ABL25999	AbL25999 Drosophil
C 38	16.2	77.1	2943	4	ABL26020	AbL26020 Drosophil
C 39	16.2	77.1	3090	5	ABV27431	Abv27431 Human pro
C 40	16.2	77.1	3090	5	ABV21611	Abv21611 Human pro
C 41	16.2	77.1	3336	8	ACF35642	Acf35642 Human nuc
C 42	16.2	77.1	3598	4	AAK79058	Aak79058 Human imm
C 43	16.2	77.1	3598	4	AAK79059	Aak79059 Human imm
C 44	16.2	77.1	3667	4	AAH16621	Aah16621 Human CDN
C 45	16.2	77.1	4423	4	AAS35053	Aas35053 DNA #3 en

## ALIGNMENTS

### RESULT 1

AAZ50445  
ID AAZ50445 standard; DNA; 21 BP.

AC AAZ50445;

DT 18-MAY-2000 (first entry)

DE EST R00504-specific primer 2.

XX PB39; human; prostate cancer; PC; chromosome 11p11.1-11.2; cancer;  
KW prostate epithelium; splicing mechanism; early diagnosis; progression;  
KW precancerous cell; metastatic potential; non-neoplastic prostate disease;  
KW expressed sequence tag; EST; PCR primer; ss.

OS Homo sapiens.

XX WO200005376-A1.

PD 03-FEB-2000.

PF 23-JUL-1999; 99WO-US016831.

XX 24-JUL-1998; 98US-0094137P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Chuquai RP, Cole KA, Liotta LA;

XX WPI; 2000-182700/16.

PT Novel gene which is dysregulated in prostate cancer useful for diagnosing cancer.

PS Claim 5; Page 16; 51pp; English.

CC The present sequence is the EST AAR00504-specific PCR primer, used for amplification of sequences contained within the EST AAR00504. It is useful to probe the gene overexpressed in prostate cancer epithelium and to analyse the differential expression of the EST. The PB39 gene that is dysregulated in prostate cancer is isolated from human pancreas cDNA library and has homology to the EST AAR00504. PB39 gene is located on chromosome 11p11.1-11.2. Abnormally high concentrations of PB39 are found in prostate tissue derived from prostate cancer (PC) epithelium. PB39 sequence is useful for detection of precancerous or cancer cells in the prostate. PB39 is useful for early diagnosis of the progression of

CC prostate cancer, especially in aggressive prostate carcinoma. It can also  
 CC distinguish PC from other non-neoplastic prostate disease. The diagnostic  
 CC method is selective and specific for various types of PC and also  
 CC facilitates identifying prostate cancer of differing aggressiveness and  
 CC metastatic potential

XX  
 SQ Sequence 21 BP; 4 A; 4 C; 7 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 3; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.49;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGCGTATCTGAAGAGTCTG 21

|||||  
 Db 1 CTGGCGTATCTGAAGAGTCTG 21

# RESULT 2

ACI19942/c

ID ACI19942 standard; DNA; 25 BP.

XX

AC ACI19942;

XX 13-OCT-2003 (first entry)

XX Human microarray DNA oligonucleotide SEQ ID NO 19933.

DE EST; ss; probe; expressed sequence tag; microarray; gene expression;

XX genetic variation; biallelic marker; polymorphism; human;

KW cross-species comparison.

XX Homo sapiens.

OS

XX US2003104410-A1.

PN

XX 05-JUN-2003.

PD

XX 15-MAR-2002; 2002US-00098263.

PF

XX 16-MAR-2001; 2001US-0276759P.

PR

XX (AFFY-) AFFYMETRIX INC.

PA

XX Mittmann MP;

PI

XX WPI; 2003-567953/53.

DR

XX New array of nucleic acid probes, useful for in situ hybridization, in

PT Southern, Northern or dot-blot hybridization to identify or detect the

PT sequence or specific mutations of any gene.

PT

XX Claim 1; SEQ ID NO 19933; 9pp; English.

PS

CC The invention discloses a microarray comprising a plurality of nucleic

CC acid probes including one of 2,018,500 fully defined sequences, or its

CC perfect match, perfect mismatch, antisense match or antisense mismatch.

CC Also disclosed is a method of gene expression analysis. The array is used

CC in monitoring gene expression levels by hybridisation to a DNA library.

CC compounds. The nucleic acid probes are specifically designed for analysis

CC of at least one target sequence. The method of analysis comprises

CC hybridising at least one or more nucleic acids to at least two or more

CC nucleic acid probes and detecting the hybridisation. The nucleic acid

CC probes are attached to a solid support. The analysis comprises monitoring

CC gene expression levels, identifying biallelic markers or polymorphisms,

CC or family members of a gene and a cross-species comparison. Each of the

CC nucleic acids further comprises a tag sequence. The array of nucleic acid

CC probes is useful in situ hybridisation, in Southern, Northern or dot-

CC blot hybridisation to identify or detect the sequence or specific

CC mutations of any gene, in mapping the 5' termini of mRNA molecules by

CC primer extensions or in screening cDNA or genomic libraries or subclones

CC for additional subclones containing segments of DNA that have been

CC isolated and previously sequenced. The sequence presented is one of the

CC nucleic acid probes incorporated in the microarray. Note: The sequence

CC data for this patent can also be obtained in electronic format directly

CC from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)

XX

SQ Sequence 25 BP; 8 A; 8 C; 4 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 8; Length 25;

Best Local Similarity 100.0%; Pred. No. 0.51;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGCGTATCTGAAGAGTCTG 21

|||||  
 Db 24 CTGGCGTATCTGAAGAGTCTG 4

# RESULT 3

ACI19943/c

ID ACI19943 standard; DNA; 25 BP.

XX

AC ACI19943;

XX

DT 13-OCT-2003 (first entry)

XX Human microarray DNA oligonucleotide SEQ ID NO 19934.

DE EST; ss; probe; expressed sequence tag; microarray; gene expression;

XX genetic variation; biallelic marker; polymorphism; human;

KW cross-species comparison.

XX Homo sapiens.

OS

XX US2003104410-A1.

PN

XX 05-JUN-2003.

PD

XX 15-MAR-2002; 2002US-00098263.

PF

XX 16-MAR-2001; 2001US-0276759P.

PR

XX (AFFY-) AFFYMETRIX INC.

PA

XX Mittmann MP;

PI

XX WPI; 2003-567953/53.

DR

XX New array of nucleic acid probes, useful for in situ hybridization, in

PT Southern, Northern or dot-blot hybridization to identify or detect the

PT sequence or specific mutations of any gene.

PT

XX Claim 1; SEQ ID NO 19934; 9pp; English.

PS

CC The invention discloses a microarray comprising a plurality of nucleic

CC acid probes including one of 2,018,500 fully defined sequences, or its

CC perfect match, perfect mismatch, antisense match or antisense mismatch.

CC Also disclosed is a method of gene expression analysis. The array is used

CC in monitoring gene expression levels by hybridisation to a DNA library,

CC compounds. The nucleic acid probes are specifically designed for analysis

CC of at least one target sequence. The method of analysis comprises

CC hybridising at least one or more nucleic acids to at least two or more

CC nucleic acid probes and detecting the hybridisation. The nucleic acid

CC probes are attached to a solid support. The analysis comprises monitoring

CC gene expression levels, identifying biallelic markers or polymorphisms,

CC or family members of a gene and a cross-species comparison. Each of the

CC nucleic acids further comprises a tag sequence. The array of nucleic acid

CC probes is useful in situ hybridisation, in Southern, Northern or dot-

CC blot hybridisation to identify or detect the sequence or specific

CC mutations of any gene, in mapping the 5' termini of mRNA molecules by

CC primer extensions or in screening cDNA or genomic libraries or subclones

CC for additional subclones containing segments of DNA that have been

CC isolated and previously sequenced. The sequence presented is one of the

CC nucleic acid probes incorporated in the microarray. Note: The sequence

CC data for this patent can also be obtained in electronic format directly



CC from USPTO at seqdata.uspto.gov/sequence.html  
 XX Sequence 25 BP; 8 A; 7 C; 5 G; 5 T; 0 U; 0 Other;  
 SQ Query Match 92.4%; Score 19.4; DB 8; Length 25;  
 Best Local Similarity 95.2%; Pred. No. 3.4;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
 DB 24 CTGGCGTATCTCAAGAGTCTG 4

RESULT 4  
 ADB54008/C  
 ID ADB54008 standard; DNA; 2455 BP.  
 XX  
 AC ADB54008;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE MSH5 genomic DNA region.  
 XX  
 KW colon cell proliferative disorder; non methylated CpG dinucleotide;  
 KW cytostatic; cancer; adenoma; carcinoma; cytosine methylation state; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003072821-A2.  
 XX  
 PD 04-SEP-2003.  
 XX  
 PF 27-FEB-2003; 2003WO-EP002035.  
 XX  
 PR 27-FEB-2002; 2002EP-00004551.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 PI Adorjan P, Burger M; Maier S, Nimmrich I, Becker E, Lesche R;  
 PI Rujan T, Schmitt A;  
 XX  
 DR WPI; 2003-731620/69.  
 XX

Detecting and differentiating between colon cell proliferative disorders associated with a gene or its regulatory regions comprises contacting a target nucleic acid in a biological sample obtained from the subject with a reagent.

Claim 46; SEQ ID NO 64; 74pp; English.

The invention relates to a novel method for detecting and differentiating between colon cell proliferative disorders associated with at least one gene or its regulatory regions. The method comprises contacting a target nucleic acid in a biological sample obtained from the subject with at least one reagent or a series of reagents, where the reagent or series of reagents, distinguishes between methylated and non methylated CpG dinucleotides within the target nucleic acid. The molecules of the invention demonstrate cytosine methylation state or single nucleotide for detecting and differentiating between colon cell proliferative disorders, including cancers such as colon adenoma and colon carcinoma. The PNA (peptide nucleic acid)-oligomers are useful as probes for determining cytosine methylation state or single nucleotide polymorphisms. The current sequence is that of the genomic DNA region of the invention. This sequence is not shown within the specification but is taken from Wipoweb.

Sequence 2455 BP; 542 A; 709 C; 592 G; 612 T; 0 U; 0 Other;  
 Query Match 84.8%; Score 17.8; DB 9; Length 2455;  
 Best Local Similarity 90.5%; Pred. No. 52;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21

Db 1888 CTGGCGTAGCTGAAGAGTCTG 1868

RESULT 5  
 ABR83555  
 ID ABR83555 standard; cDNA; 28772 BP.  
 XX  
 AC ABR83555;  
 XX  
 DT 14-AUG-2002 (first entry)  
 XX  
 DE Human cDNA differentially expressed in granulocytic cells #126.  
 XX  
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
 KW viral infection; parasitic infection; protozoal infection;  
 KW fungal infection; sterile inflammatory disease; psoriasis;  
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
 KW adult respiratory distress syndrome; inflammatory bowel disease;  
 KW Crohn's disease; ulcerative colitis; periodontal disease;  
 KW granulocyte activation; chronic inflammation; allergy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200228999-A2.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 03-OCT-2001; 2001WO-US030821.  
 XX  
 PR 03-OCT-2000; 2000US-0237189P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 XX  
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
 XX  
 DR WPI; 2002-435328/46.  
 XX

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity.

Claim 1; SEQ ID NO 126; 114pp; English.

The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial



Query Match 78.1%; Score 16.4; DB 2; Length 3357;  
Best Local Similarity 94.4%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 GGCCTATCTGAACAGTCT 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1959 GGCCTATCTGAACAGTCT 1976

RESULT 8  
AAI99253  
ID AAI99253 standard; DNA; 13435 BP.  
XX  
AC AAI99253;  
XX  
DT 07-JAN-2002 (first entry)  
XX  
DE Human excretory related polynucleotide SEQ ID NO 1017.  
XX  
KW Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antisickling; antianemic; antiarthritic; cancer;  
KW antineumatic; hepatotropic; cerebroprotective; antinflamatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;  
KW excretory system; ds.  
XX  
OS Homo sapiens.  
XX  
PN WC200153113-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US0001323.  
XX  
31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.

```
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465569/50.
XX
XX Isolated nucleic acid molecule encoding excretory system antigen is used
XX in preventing, treating or ameliorating a medical condition.
XX
XX Example 2; SEQ ID NO 1017; 574pp + Sequence Listing; English.
XX
XX The invention relates to novel excretory system related human
XX polynucleotides (AA198567-AA199503) and the encoded proteins (AAM99594-
XX AAM99913) useful for preventing, treating or ameliorating medical
XX conditions e.g. by protein or gene therapy, especially disorders related
XX to the excretory system. The genes are isolated from a range of human
XX tissues disclosed in the specification. The nucleic acids, proteins, and
XX antibodies and (anti)agonists are useful in the diagnosis, treatment and
XX prevention of: (a) cancer, e.g. breast and ovarian cancer and other
XX cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal
XX tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's
XX disease, allergies, autoimmune haemolytic anaemia, autoimmune
XX rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders
XX such as myocardial ischaemia; (d) wound healing; (e) neurological
XX diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases
XX such as viral, bacterial, fungal and parasitic infections. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 13435 BP; 3509 A; 2951 C; 3221 G; 3754 T; 0 U; 0 Other;
XX
XX Query Match
XX Best Local Similarity 78.1%; Score 16.4; DB 4; Length 13435;
XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 CTGGCGTATCTGAAGT 18
XX ||||| ||||| |||||
XX Db 8309 CTGGCTTATCTGAAGT 8326
XX
XX RESULT 9
XX AAL04187
XX ID AAL04187 standard; DNA; 13435 BP.
```

```
XX AAL04187;
XX
XX 21-NOV-2001 (first entry)
XX
XX Human reproductive system related antigen DNA SEQ ID NO: 6875.
XX
XX Human; reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX WO200155320-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001339.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226686P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
```

```
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-02344223P.
PR 21-SEP-2000; 2000US-0234423P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0246417P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
XX used in preventing, treating or ameliorating a medical condition.
XX Disclosure; SEQ ID NO 6875; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX CC including cancer. The present sequence is a genomic sequence encoding a
XX protein of the invention
XX
XX SQ Sequence 13435 BP; 3509 A; 2951 C; 3221 G; 3754 T; 0 U; 0 Other;
XX
Query Match 78.1%; Score 16.4; DB 4; Length 13435;
Best Local Similarity 94.4%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTGGCGTATCTGAAGAGT 18
DB 8309 CTGGCTTATCTGAAGAGT 8326
RESULT 10
AAI63603
ID AAI63603 standard; DNA; 13435 BP.
XX
XX AAI63603;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human kidney related polynucleotide SEQ ID NO 918.
KW Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; antiallergic; hepatotropic; antidiabetic;
KW antinflammatory; antiulcer; vulnery; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
XX
XX Homo sapiens.
XX
XX WO200155323-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001343.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
```

PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0228273P.  
 PR 22-AUG-2000; 2000US-0228681P.  
 PR 22-AUG-2000; 2000US-0228688P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0228287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0233080P.  
 PR 08-SEP-2000; 2000US-0233081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-023400P.  
 PR 14-SEP-2000; 2000US-023401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-023423P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 PR Rosen CA, Barash SC, Ruben SM;  
 PR WPI; 2001-488784/53.  
 PR New isolated nucleic acids and polypeptides, useful for diagnosing,  
 PR treating and/or preventing human diseases and disorders.  
 PR Disclosure; SEQ ID NO 918; 564pp + Sequence Listing; English.  
 PR The invention relates to novel kidney related polynucleotides (AA162971-  
 PR AA163793) and the encoded polypeptides (AA42417-AA42691) collectively  
 PR known as kidney antigens and the use of such kidney antigens for  
 PR detecting disorders of the kidney, especially kidney cancer and kidney  
 PR cancer metastases. The polynucleotides and proteins are also useful for  
 PR preventing, treating or ameliorating medical conditions e.g. by protein  
 PR or gene therapy. The genes are isolated from a range of human tissues  
 PR disclosed in the specification. The nucleic acids, proteins, antibodies

CC and (ant)agonists are useful in the diagnosis, treatment and prevention  
 CC of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the  
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as  
 CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.  
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,  
 CC bacterial, fungal and parasitic infections. Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 13435 BP; 3509 A; 2951 C; 3221 G; 3754 T; 0 U; 0 Other;  
 Query Match 78.1%; Score 16.4; DB 5; Length 13435;  
 Best Local Similarity 94.4%; Pred. No. 3.7e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CTGGCGTATCTGAAGAGT 18  
 ||||| ||||| ||||| |||||  
 Db 8309 CTGGCTATCTGAAGAGT 8326

RESULT 11  
 ACC69924  
 ID ACC69924 standard; DNA; 36 BP.

XX ACC69924;

XX 28-JUL-2003 (first entry)

XX Human TRAP oligonucleotide probe SEQ ID NO:5.

XX Bone growth; osteoblast; osteoclast; TRAP; TRIP-1; glypican 4; GPC4;  
 KW transforming growth factor beta receptor interacting protein 1;  
 KW tartrate resistant acid phosphatase; TGF-beta; osteopathic; litholytic;  
 KW nephrotropic; antidiabetic; cytostatic; antitumour; antirheumatic;  
 KW antiarthritic; bone formation regulation; osteoclast lacunae; diabetes;  
 KW osteoporosis; renal osteodystrophy; renal stones; hyperparathyroidism;  
 KW hyperthyroidism; hypercalcaemia; Fanconi's syndrome; sarcoidosis;  
 KW osteomalacia; nutritional rickets; hypervitaminosis; paget's disease;  
 KW osteopetrosis; skeletal tumour; rheumatoid arthritis; osteoarthritis;  
 KW osteogenesis imperfecta; chondrodystrophy; sclerosing bone dysplasia;  
 KW probe; ss.

XX Homo sapiens.  
 OS Synthetic.

XX WO2003024408-A2.

XX 27-MAR-2003.

XX 20-SEP-2002; 2002WO-US030093.

XX 20-SEP-2001; 2001US-0323987P.

XX (UYRP ) UNIV ROCHESTER.

XX Puzas JE;

XX WPI; 2003-354553/33.

XX Novel composition useful for producing transgenic animals, contains a  
 PT peptide that binds to tartrate resistant acid phosphatase.

XX Example 1; Page 96; 196pp; English.

XX The present invention describes an isolated composition (I), comprising a  
 CC peptide (II) that binds tartrate resistant acid phosphatase (TRAP). Also  
 CC described is a method for producing a peptide that binds TRAP, by  
 CC expressing the peptide from a nucleic acid or chemically synthesising the

CC peptide (where the peptide comprises a sequence of ABR44759 to ABR44776.  
 CC (I) has osteopathic, litholytic, nephrotropic, antidiabetic, cytostatic,  
 CC antitumour, antirheumatic and antiarthritic activities, and can be used  
 CC to regulate bone formation and modulate osteoblast binding to osteoclast  
 CC lacunae. (I) can be used for producing a cell comprising a composition  
 CC that binds to TRAP. (I) can also be used for regulating bone formation,  
 CC where (I) binds glypican (GPC)-4 or transforming growth factor (TGF)-beta  
 CC receptor interacting protein (TRIP). The method involves decreasing bone  
 CC formation by inhibiting osteoblast binding to osteoclast lacunae, or  
 CC increasing bone formation by increasing osteoblast binding to osteoclast  
 CC lacunae or by increasing osteoblast differentiation. The method is useful  
 CC for treating osteoporosis, renal osteodystrophy, renal stones, diabetes,  
 CC hyperparathyroidism, hyperthyroidism, hypercalcaemia, Fanconi's syndrome,  
 CC sarcoidosis, osteomalacia, nutritional rickets, hypervitaminosis A and D,  
 CC paget's disease, osteopetrosis, skeletal tumours, rheumatoid and osteo-  
 CC arthritis, osteogenesis imperfecta, chondrodystrophies or sclerosing bone  
 CC dysplasias. The present sequence represents a TRAP oligonucleotide probe,  
 CC which is used in an example from the present invention

XX Sequence 36 BP; 5 A; 8 C; 13 G; 10 T; 0 U; 0 Other;

Query Match 77.1%; Score 16.2; DB 7; Length 36;  
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CTGGCGTATCTGAAGAGTCTG 21

||| ||||| ||||| ||||| |||||  
 Db 7 CTTCGTATCTGAAGGCTCTG 27

RESULT 12

ACC69943

ID ACC69943 standard; DNA; 36 BP.

XX ACC69943;

XX 28-JUL-2003 (first entry)

XX Human TRAP degenerate oligonucleotide SEQ ID NO:46.

XX Bone growth; osteoblast; osteoclast; TRAP; TRIP-1; glypican 4; GPC4;  
 KW transforming growth factor beta receptor interacting protein 1;  
 KW tartrate resistant acid phosphatase; TGF-beta; osteopathic; litholytic;  
 KW nephrotropic; antidiabetic; cytostatic; antitumour; antirheumatic;  
 KW antiarthritic; bone formation regulation; osteoclast lacunae; diabetes;  
 KW osteoporosis; renal osteodystrophy; renal stones; hyperparathyroidism;  
 KW hyperthyroidism; hypercalcaemia; Fanconi's syndrome; sarcoidosis;  
 KW osteomalacia; nutritional rickets; hypervitaminosis; paget's disease;  
 KW osteopetrosis; skeletal tumour; rheumatoid arthritis; osteoarthritis;  
 KW osteogenesis imperfecta; chondrodystrophy; sclerosing bone dysplasia;  
 KW chromosome 19; enzyme; gene; ss.

XX Homo sapiens.

XX WO2003024408-A2.

XX 27-MAR-2003.

XX 20-SEP-2002; 2002WO-US030093.

XX 20-SEP-2001; 2001US-0323987P.

XX (UYRP ) UNIV ROCHESTER.

XX Puzas JE;

XX WPI; 2003-354553/33.

XX Novel composition useful for producing transgenic animals, contains a  
 PT peptide that binds to tartrate resistant acid phosphatase.

XX Disclosure; Page 137; 196pp; English.

CC The present invention describes an isolated composition (I), comprising a  
 CC peptide (II) that binds tartrate resistant acid phosphatase (TRAP). Also  
 CC described is a method for producing a peptide that binds TRAP, by  
 CC expressing the peptide from a nucleic acid or chemically synthesizing the  
 CC peptide (where the peptide comprises a sequence of ABR44759 to ABR44776.  
 CC (I) has osteopathic, litholytic, nephrotropic, antidiabetic, cytostatic,  
 CC antitumour, antirheumatic and antiarthritic activities, and can be used  
 CC to regulate bone formation and modulate osteoblast binding to osteoclast  
 CC lacunae. (I) can be used for producing a cell comprising a composition  
 CC that binds to TRAP. (I) can also be used for regulating bone formation,  
 CC where (I) binds glypican (GPC)-4 or transforming growth factor (TGF)-beta  
 CC receptor interacting protein (TRIP). The method involves decreasing bone  
 CC formation by inhibiting osteoblast binding to osteoclast lacunae, or  
 CC increasing bone formation by increasing osteoblast binding to osteoclast  
 CC lacunae or by increasing osteoblast differentiation. The method is useful  
 CC for treating osteoporosis, renal osteodystrophy, renal stones, diabetes,  
 CC hyperparathyroidism, hyperthyroidism, hypercalcaemia, Fanconi's syndrome,  
 CC sarcoidosis, osteomalacia, nutritional rickets, hypervitaminosis A and D,  
 CC paget's Disease, osteopetrosis, skeletal tumours, rheumatoid and osteo-  
 CC arthritis, osteogenesis imperfecta, chondrodystrophies or sclerosing bone  
 CC dysplasias. The present sequence represents a human tartrate resistant  
 CC acid phosphatase (TRAP) degenerate oligonucleotide, which is given in the  
 CC exemplification of the present invention. Human TRAP is located on  
 CC chromosome 19

SQ Sequence 36 BP; 7 A; 8 C; 10 G; 11 T; 0 U; 0 Other;

Query Match 77.1%; Score 16.2; DB 7; Length 36;

Best Local Similarity 85.7%; Pred. No. 1.6e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21

DB 7 CTTTCGTATCTGAAGGCTG 27

RESULT 13

ACC69942

ID ACC69942 standard; DNA; 36 BP.

XX ACC69942;

AC ACC69942;

DT 28-JUL-2003 (first entry)

DE Human TRAP degenerate oligonucleotide SEQ ID NO:45.

XX Bone growth; osteoblast; osteoclast; TRAP; TRIP-1; glypican 4; GPC4;  
 KW transforming growth factor beta receptor interacting protein 1;  
 KW tartrate resistant acid phosphatase; TGF-beta; osteopathic; litholytic;  
 KW nephrotropic; antidiabetic; cytostatic; antitumour; antirheumatic;  
 KW antiarthritic; bone formation regulation; osteoclast lacunae; diabetes;  
 KW osteoporosis; renal osteodystrophy; renal stones; hyperparathyroidism;  
 KW hyperthyroidism; hypercalcaemia; Fanconi's syndrome; sarcoidosis;  
 KW osteomalacia; nutritional rickets; hypervitaminosis; paget's Disease;  
 KW osteopetrosis; skeletal tumour; rheumatoid arthritis; osteoarthritis;  
 KW osteogenesis imperfecta; chondrodystrophy; sclerosing bone dysplasia;  
 KW chromosome 19; enzyme; gene; ss.

XX Homo sapiens.

OS WO2003024408-A2.

XX 27-MAR-2003.

XX 20-SEP-2002; 2002WO-US030093.

XX 20-SEP-2001; 2001US-0323987P.

XX (UYRP ) UNIV ROCHESTER.

XX Puzas JE;

XX WPI; 2003-354553/33.

XX

PT Novel composition useful for producing transgenic animals, contains a  
 PT peptide that binds to tartrate resistant acid phosphatase.

XX Disclosure; Page 137; 196pp; English.

XX

CC The present invention describes an isolated composition (I), comprising a  
 CC peptide (II) that binds tartrate resistant acid phosphatase (TRAP). Also  
 CC described is a method for producing a peptide that binds TRAP, by  
 CC expressing the peptide from a nucleic acid or chemically synthesizing the  
 CC peptide (where the peptide comprises a sequence of ABR44759 to ABR44776.  
 CC (I) has osteopathic, litholytic, nephrotropic, antidiabetic, cytostatic,  
 CC antitumour, antirheumatic and antiarthritic activities, and can be used  
 CC to regulate bone formation and modulate osteoblast binding to osteoclast  
 CC lacunae. (I) can be used for producing a cell comprising a composition  
 CC that binds to TRAP. (I) can also be used for regulating bone formation,  
 CC where (I) binds glypican (GPC)-4 or transforming growth factor (TGF)-beta  
 CC receptor interacting protein (TRIP). The method involves decreasing bone  
 CC formation by inhibiting osteoblast binding to osteoclast lacunae, or  
 CC increasing bone formation by increasing osteoblast binding to osteoclast  
 CC lacunae or by increasing osteoblast differentiation. The method is useful  
 CC for treating osteoporosis, renal osteodystrophy, renal stones, diabetes,  
 CC hyperparathyroidism, hyperthyroidism, hypercalcaemia, Fanconi's syndrome,  
 CC sarcoidosis, osteomalacia, nutritional rickets, hypervitaminosis A and D,  
 CC paget's Disease, osteopetrosis, skeletal tumours, rheumatoid and osteo-  
 CC arthritis, osteogenesis imperfecta, chondrodystrophies or sclerosing bone  
 CC dysplasias. The present sequence represents a human tartrate resistant  
 CC acid phosphatase (TRAP) degenerate oligonucleotide, which is given in the  
 CC exemplification of the present invention. Human TRAP is located on  
 CC chromosome 19

SQ Sequence 36 BP; 8 A; 8 C; 10 G; 10 T; 0 U; 0 Other;

Query Match 77.1%; Score 16.2; DB 7; Length 36;

Best Local Similarity 85.7%; Pred. No. 1.6e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21

DB 7 CTTTCGTATCTGAAGGCTG 27

RESULT 14

ACC69941

ID ACC69941 standard; DNA; 36 BP.

XX ACC69941;

AC ACC69941;

DT 28-JUL-2003 (first entry)

DE Human TRAP degenerate oligonucleotide SEQ ID NO:43.

XX Bone growth; osteoblast; osteoclast; TRAP; TRIP-1; glypican 4; GPC4;  
 KW transforming growth factor beta receptor interacting protein 1;  
 KW tartrate resistant acid phosphatase; TGF-beta; osteopathic; litholytic;  
 KW nephrotropic; antidiabetic; cytostatic; antitumour; antirheumatic;  
 KW antiarthritic; bone formation regulation; osteoclast lacunae; diabetes;  
 KW osteoporosis; renal osteodystrophy; renal stones; hyperparathyroidism;  
 KW hyperthyroidism; hypercalcaemia; Fanconi's syndrome; sarcoidosis;  
 KW osteomalacia; nutritional rickets; hypervitaminosis; paget's Disease;  
 KW osteopetrosis; skeletal tumour; rheumatoid arthritis; osteoarthritis;  
 KW osteogenesis imperfecta; chondrodystrophy; sclerosing bone dysplasia;  
 KW chromosome 19; enzyme; gene; ss.

XX Homo sapiens.

OS WO2003024408-A2.

XX 27-MAR-2003.

XX 20-SEP-2002; 2002WO-US030093.

XX 20-SEP-2001; 2001US-0323987P.



XX PA (UYRP ) UNIV ROCHESTER.  
XX PI Puzas JE;  
XX PI WPI; 2003-354553/33.  
XX PT Novel composition useful for producing transgenic animals, contains a  
XX PT peptide that binds to tartrate resistant acid phosphatase.  
XX PS Disclosure; Page 136; 196pp; English.  
XX CC The present invention describes an isolated composition (I), comprising a  
XX CC peptide (II) that binds tartrate resistant acid phosphatase (TRAP). Also  
XX CC described is a method for producing a peptide that binds TRAP, by  
XX CC expressing the peptide from a nucleic acid or chemically synthesizing the  
XX CC peptide (where the peptide comprises a sequence of ABR44759 to ABR44776.  
XX CC (I) has osteopontin, litholytic, nephrotropic, antidiabetic, cytostatic,  
XX CC antitumor, antirheumatic and antiarthritic activities, and can be used  
XX CC to regulate bone formation and modulate osteoblast binding to osteoclast  
XX CC lacunae. (I) can be used for producing a cell comprising a composition  
XX CC that binds to TRAP. (I) can also be used for regulating bone formation,  
XX CC where (I) binds glypican (GPC)-4 or transforming growth factor (TGF)-beta  
XX CC receptor interacting protein (TRIP). The method involves decreasing bone  
XX CC formation by inhibiting osteoblast binding to osteoclast lacunae, or  
XX CC increasing bone formation by increasing osteoblast binding to osteoclast  
XX CC lacunae or by increasing osteoblast differentiation. The method is useful  
XX CC for treating osteoporosis, renal osteodystrophy, renal stones, diabetes,  
XX CC hyperparathyroidism, hyperthyroidism, hypercalcaemia, Fanconi's syndrome,  
XX CC sarcoidosis, osteomalacia, nutritional rickets, hypervitaminosis A and D,  
XX CC Paget's Disease, osteopetrosis, skeletal tumours, rheumatoid and osteo-  
XX CC arthritis, osteogenesis imperfecta, chondrodystrophies or sclerosing bone  
XX CC dysplasias. The present sequence represents a human tartrate resistant  
XX CC acid phosphatase (TRAP) degenerate oligonucleotide, which is given in the  
XX CC exemplification of the present invention. Human TRAP is located on  
XX CC chromosome 19  
XX SQ Sequence 36 BP; 6 A; 8 C; 12 G; 10 T; 0 U; 0 Other;  
Query Match 77.1%; Score 16.2; DB 7; Length 36;  
Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CTGGCGTATCTGAAGAGTCTG 21  
Db 7 CTTTCGTATCTGAAGGCTG 27  
RESULT 15  
ID ABS69156  
XX ABS69156 standard; DNA; 230 BP.  
XX AC ABS69156;  
XX DT 21-NOV-2002 (first entry)  
XX DE Novel murine polynucleotide isolated using gene trap technology #219.  
XX KW Mouse; gene trapped sequence; GTS; functional genomic analysis;  
XX KW phage display system; gene chip; temporal gene expression;  
XX KW tissue specific gene expression; antisense inhibition; gene targeting;  
XX KW development disorder; cell differentiation disorder; aging; cancer;  
XX KW autoimmune disease; lupus; inflammatory disorder; skin disorder;  
XX KW degenerative disorder; ds.  
XX OS Mus musculus.  
XX PN US2002102543-A1.  
XX PD 01-AUG-2002.  
XX PF 30-NOV-2000; 2000US-00728445.

PR 01-DEC-1999; 99US-0168358P.  
XX (FRIE/) FRIEDRICH G.  
XX PA (ZAMB/) ZAMBROWICZ B.  
XX PA (SAND/) SANDS A T.  
XX PI Friedrich G, Zambrowicz B, Sands AT;  
XX WPI; 2002-690598/74.  
XX PT Novel murine polynucleotides that individually identify novel genes into  
XX PT which a retroviral gene trap vector has integrated, useful in genomic  
XX PT analysis and in discovery, development of therapeutic and diagnostic  
XX PT agents.  
XX PS Claim 1; Page 89; 296pp; English.  
XX CC The invention describes an isolated murine polynucleotide (I) comprising  
XX CC a contiguous stretch of at least 60 nucleotides of one of 265-677  
XX CC nucleotide 891 OMIBANK gene trapped sequences (GTSs) (S). Given in the  
XX CC specification. The novel genes and cells are useful in functional genomic  
XX CC analysis and in the discovery and development of new therapeutic and  
XX CC diagnostic agents and methods. (I) is useful for identifying the coding  
XX CC regions of the murine genome, to isolate cDNAs, genomic clones, or full-  
XX CC length genes/polynucleotides or homologues, heterologues, paralogues, or  
XX CC orthologues that are capable of hybridising to one or more of the GTSs  
XX CC under stringent conditions. (I) can be incorporated into a phage display  
XX CC system that can be used to screen for proteins, or other ligands, that  
XX CC are capable of binding an amino acid sequence encoded by an  
XX CC oligonucleotide or polynucleotide sequence in at least one of the TS  
XX CC sequences. (I) is useful in addressable arrays, such as gene chips, to  
XX CC identify and characterise temporal and tissue specific gene expression,  
XX CC to identify the gene of interest from many sources and for genetic  
XX CC manipulations such as antisense inhibition and gene targeting. Decreasing  
XX CC the level of expression of (I) and/or down regulating the activity of  
XX CC peptides or proteins encoded by (I) is useful for treating development  
XX CC and cell differentiation disorders, aging, cancer, autoimmune disease,  
XX CC lupus, inflammatory disorders, skin disorders and degenerative disorders.  
XX CC This sequence represents a murine cDNA isolated using gene trap  
XX CC technology  
XX SQ Sequence 230 BP; 58 A; 55 C; 53 G; 62 T; 0 U; 2 Other;  
Query Match 77.1%; Score 16.2; DB 6; Length 230;  
Best Local Similarity 85.7%; Pred. No. 2.3e+02;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CTGGCGTATCTGAAGAGTCTG 21  
Db 184 CTGGTGTGTCTGAAGAGACTG 204

Search completed: April 10, 2004, 20:05:49  
Job time : 10.8079 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 14:52:03 ; Search time 1.94265 Seconds  
(without alignments)  
5998.994 Million cell updates/sec

Title: US-09-743-825-8

Perfect score: 21

Sequence: 1 ctggcgatctgaagagtctg 21

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/prodata/2/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.2	77.1	597	4	US-08-858-207A-94
2	16.2	77.1	613	4	US-09-023-655-350
3	16.2	77.1	2876	4	US-09-976-594-246
4	16.2	77.1	11831	4	US-08-961-527-65
5	16	76.2	714	4	US-09-172-952-9
6	16	76.2	9334	4	US-09-172-952-19
7	15.4	73.3	720	3	US-09-479-309-3
8	15.4	73.3	720	4	US-09-627-393-3
9	15.4	73.3	1654	4	US-09-620-312B-581
10	15.2	72.4	30	2	US-08-859-998-456
11	15.2	72.4	30	4	US-09-225-928-456
12	15.2	72.4	30	4	US-09-225-201B-456
13	15.2	72.4	415	2	US-08-702-105A-24
14	15.2	72.4	415	3	US-08-702-105A-26
15	15.2	72.4	415	3	US-08-702-110A-24
16	15.2	72.4	415	3	US-08-702-110A-26
17	15.2	72.4	415	3	US-09-325-571-24
18	15.2	72.4	415	3	US-09-325-571-26
19	15.2	72.4	415	4	US-09-848-585-24
20	15.2	72.4	415	4	US-09-848-585-26
21	15.2	72.4	536	6	5208218-2
22	15.2	72.4	585	2	US-08-702-105A-28
23	15.2	72.4	585	2	US-08-702-105A-30
24	15.2	72.4	585	3	US-08-702-110A-28
25	15.2	72.4	585	3	US-08-702-110A-30
26	15.2	72.4	585	3	US-09-325-571-28
27	15.2	72.4	585	3	US-09-325-571-30

28	15.2	72.4	585	4	US-09-848-585-28
29	15.2	72.4	585	4	US-09-848-585-30
30	15.2	72.4	612	4	US-09-540-236-1227
31	15.2	72.4	786	1	US-08-345-756-1
32	15.2	72.4	786	1	US-08-625-198-1
33	15.2	72.4	1035	4	US-09-489-039A-916
34	15.2	72.4	1229	4	US-08-836-047-2
35	15.2	72.4	1534	4	US-09-489-847-110
36	15.2	72.4	1674	4	US-09-489-847-18
37	15.2	72.4	2039	1	US-08-345-756-5
38	15.2	72.4	2039	1	US-08-625-198-5
39	15.2	72.4	2276	3	US-09-172-841-56
40	15.2	72.4	2276	4	US-08-951-621-56
41	15.2	72.4	89047	4	US-09-596-003-34
42	15	71.4	970	3	US-08-545-809A-40
43	14.8	70.5	61	3	US-09-030-613-26
44	14.8	70.5	61	4	US-09-451-903-26
45	14.8	70.5	360	4	US-09-221-017B-298

ALIGNMENTS

RESULT 1  
US-08-858-207A-94  
; Sequence 94, Application US/08858207A  
; Patent No. 6348328  
; GENERAL INFORMATION:  
; APPLICANT: Black, Michael  
; APPLICANT: Hodgson, John  
; APPLICANT: Knowles, David  
; APPLICANT: Nicholas, Richard  
; APPLICANT: Stodola, Robert  
; TITLE OF INVENTION: No. 6348328el Compounds  
; NUMBER OF SEQUENCES: 552  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/858,207A  
; FILING DATE: 09-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/017670  
; FILING DATE: 14-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gimmi, Edward R  
; REGISTRATION NUMBER: 38,891  
; REFERENCE/DOCKET NUMBER: P50475  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-4478  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 94:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 597 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-858-207A-94

Query Match 77.1%; Score 16.2; DB 4; Length 597;  
Best Local Similarity 85.7%; Pred. No. 38;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
|||||  
Db 352 CTGGTGTATCTGAAGAGATTG 372

## RESULT 2

US-09-023-655-350  
; Sequence 350, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HERewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCES/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 350:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 613 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PLACNOT02  
; CLONE: 1445507  
US-09-023-655-350

Query Match 77.1%; Score 16.2; DB 4; Length 613;  
Best Local Similarity 85.7%; Pred. No. 38;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
|||||  
Db 147 CTGGAGTATCCGAGAGTCTG 167

## RESULT 3

US-09-976-594-246  
; Sequence 246, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US

; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 246  
; LENGTH: 2876  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 1445507CBI  
US-09-976-594-246

Query Match 77.1%; Score 16.2; DB 4; Length 2876;  
Best Local Similarity 85.7%; Pred. No. 47;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
|||||  
Db 1021 CTGGAGTATCCGAGAGTCTG 1041

## RESULT 4

US-08-961-527-65  
; Sequence 65, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCES/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11831 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-65

Query Match 77.1%; Score 16.2; DB 4; Length 11831;  
Best Local Similarity 85.7%; Pred. No. 58;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGCGTATCTGAAGAGTCTG 21  
|||||  
Db 4131 CTGGTGTATCTGAAGAGATTG 4151

```
RESULT 5
US-09-172-952-9/c
; Sequence 9, Application US/09172952
; Patent No. 6368793
; GENERAL INFORMATION:
; APPLICANT: Hoch, James
; TITLE OF INVENTION: METABOLIC SELECTION METHODS
; FILE REFERENCE: 234/191
; CURRENT APPLICATION NUMBER: US/09/172,952
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Yia S
US-09-172-952-9

Query Match          76.2%; Score 16; DB 4; Length 714;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 TGGCGTATCTGAAGAG 17
        |||||
Db      234 TGGCGTATCTGAAGAG 219

RESULT 6
US-09-172-952-19/c
; Sequence 19, Application US/09172952
; Patent No. 6368793
; GENERAL INFORMATION:
; APPLICANT: Hoch, James
; TITLE OF INVENTION: METABOLIC SELECTION METHODS
; FILE REFERENCE: 234/191
; CURRENT APPLICATION NUMBER: US/09/172,952
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 9334
; TYPE: DNA
; ORGANISM: Yia
US-09-172-952-19

Query Match          76.2%; Score 16; DB 4; Length 9334;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 TGGCGTATCTGAAGAG 17
        |||||
Db      8460 TGGCGTATCTGAAGAG 8445

RESULT 7
US-09-479-309-3
; Sequence 3, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTSD0630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 720
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-3

Query Match          73.3%; Score 15.4; DB 3; Length 720;
Best Local Similarity 94.1%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 TGGCGTATCTGAAGAGT 18
        |||||
Db      2 TGGCGTCTCTGAAGAGT 18

RESULT 8
US-09-627-393-3
; Sequence 3, Application US/09627393
; Patent No. 6534267
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTSD0630
; CURRENT APPLICATION NUMBER: US/09/627,393
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/479,309
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-627-393-3

Query Match          73.3%; Score 15.4; DB 4; Length 720;
Best Local Similarity 94.1%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 TGGCGTATCTGAAGAGT 18
        |||||
Db      2 TGGCGTCTCTGAAGAGT 18

RESULT 9
US-09-620-312D-581/c
; Sequence 581, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Felyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NO. 6569662el Nucleic Acids and
; POLYPEPTIDES
```

FILE REFERENCE: 784CIP28  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pt\_FL\_genes Version 1.0  
SEQ ID NO 581  
LENGTH: 1654  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (496)..(1302)  
US-09-620-312D-581

Query Match 73.3%; Score 15.4; DB 4; Length 1654;  
Best Local Similarity 94.1%; Pred. No. 1,1e-02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TGGCGTATCTGAAGAGT 18  
||| |||||  
Db 1592 TGGCTATCTGAAGAGT 1576

RESULT 10  
US-08-859-998-456/c  
Sequence 456, Application US/0885998  
Patent No. 5994076  
GENERAL INFORMATION:  
APPLICANT: Chenchik, Alex  
APPLICANT: Jokhadze, George  
APPLICANT: Bibilashvili, Robert  
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL  
NUMBER OF SEQUENCES: 1375  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 2200 Sand Hill Road, Suite 100  
CITY: Menlo Park  
STATE: CA  
COUNTRY: US  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/859,998  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Field, Bret E.  
REGISTRATION NUMBER: 37,620  
REFERENCE/DOCKET NUMBER: 09096/002001  
TELEPHONE: 415-322-5070  
TELEFAX: 415-854-0875  
INFORMATION FOR SEQ ID NO: 456:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
OTHER INFORMATION: oligonucleotide primer

US-08-859-998-456

Query Match 72.4%; Score 15.2; DB 2; Length 30;  
Best Local Similarity 85.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGAAGAGTCT 20  
||| |||||  
Db 24 CTGACATTTCTGAAGAGTCT 5

RESULT 11  
US-09-925-928-456/c  
Sequence 456, Application US/09225928  
Patent No. 6352829  
GENERAL INFORMATION:  
APPLICANT: Chenchik, Alex  
APPLICANT: Jokhadze, George  
APPLICANT: Bibilashvili, Robert  
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL  
NUMBER OF SEQUENCES: 1375  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 2200 Sand Hill Road, Suite 100  
CITY: Menlo Park  
STATE: CA  
COUNTRY: US  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/225,928  
FILING DATE: 05-Jan-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/859,998  
FILING DATE: 21-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Field, Bret E.  
REGISTRATION NUMBER: 37,620  
REFERENCE/DOCKET NUMBER: 09096/002001  
TELEPHONE: 415-322-5070  
TELEFAX: 415-854-0875  
INFORMATION FOR SEQ ID NO: 456:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
OTHER INFORMATION: oligonucleotide primer  
SEQUENCE DESCRIPTION: SEQ ID NO: 456:

Query Match 72.4%; Score 15.2; DB 4; Length 30;  
Best Local Similarity 85.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTGGCGTATCTGAAGAGTCT 20  
||| |||||  
Db 24 CTGACATTTCTGAAGAGTCT 5

RESULT 12  
US-09-225-201B-456/c  
Sequence 456, Application US/09225201B  
Patent No. 6489455

GENERAL INFORMATION:  
APPLICANT: Chenchik, Alex  
Bibilashvili, Robert  
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL EXPRESSION  
NUMBER OF SEQUENCES: 1375  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 2200 Sand Hill Road, Suite 100  
CITY: Menlo Park  
STATE: CA  
COUNTRY: US  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/225,201B  
FILING DATE: 05-Jan-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/859,998  
FILING DATE: 21-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Field, Bret E.  
REGISTRATION NUMBER: 37,620  
REFERENCE/DOCKET NUMBER: 09096/002001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5070  
TELEFAX: 415-854-0875  
INFORMATION FOR SEQ ID NO: 456:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
OTHER INFORMATION: oligonucleotide primer  
SEQUENCE DESCRIPTION: SEQ ID NO: 456:  
US-09-225-201B-456  
Query Match 72.4%; Score 15.2; DB 4; Length 30;  
Best Local Similarity 85.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CTGGCGTATCTGAAGAGTCT 20  
||| |  
Db 24 CTGACATTTCTGAAGAGTCT 5

RESULT 13  
US-08-702-105A-24  
Sequence 24, Application US/08702105A  
Patent No. 5908839  
GENERAL INFORMATION:  
APPLICANT: Levitt, Roy C.  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, U. Prasad  
TITLE OF INVENTION: Asthma Associated Factors As Targets For  
TITLE OF INVENTION: Treating Atopic Allergies Including Asthma And Related  
TITLE OF INVENTION: Disorders  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner L.L.P.  
STREET: 1300 I Street N.W., Suite 700  
CITY: Washington  
STATE: D.C.

COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/702,105A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/874,503  
FILING DATE: 13-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32984  
REFERENCE/DOCKET NUMBER: 05387.0056-01000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 415 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-702-105A-24  
Query Match 72.4%; Score 15.2; DB 2; Length 415;  
Best Local Similarity 85.0%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CTGGCGTATCTGAAGAGTCT 20  
||| |  
Db 317 CTGACATTTCTGAAGAGTCT 336

RESULT 14  
US-08-702-105A-26  
Sequence 26, Application US/08702105A  
Patent No. 5908839  
GENERAL INFORMATION:  
APPLICANT: Levitt, Roy C.  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, U. Prasad  
APPLICANT: Nicolaides, Nicholas C.  
TITLE OF INVENTION: Asthma Associated Factors As Targets For  
TITLE OF INVENTION: Treating Atopic Allergies Including Asthma And Related  
TITLE OF INVENTION: Disorders  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner L.L.P.  
STREET: 1300 I Street N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/702,105A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/874,503  
FILING DATE: 13-JUN-1997  
ATTORNEY/AGENT INFORMATION:

